Fifth Annual Workshop on Space Operations Applications and Research (SOAR '91)

S O A R • 9 1

Proceedings of a workshop held in
Houston, Texas
July 9–11, 1991

NASA
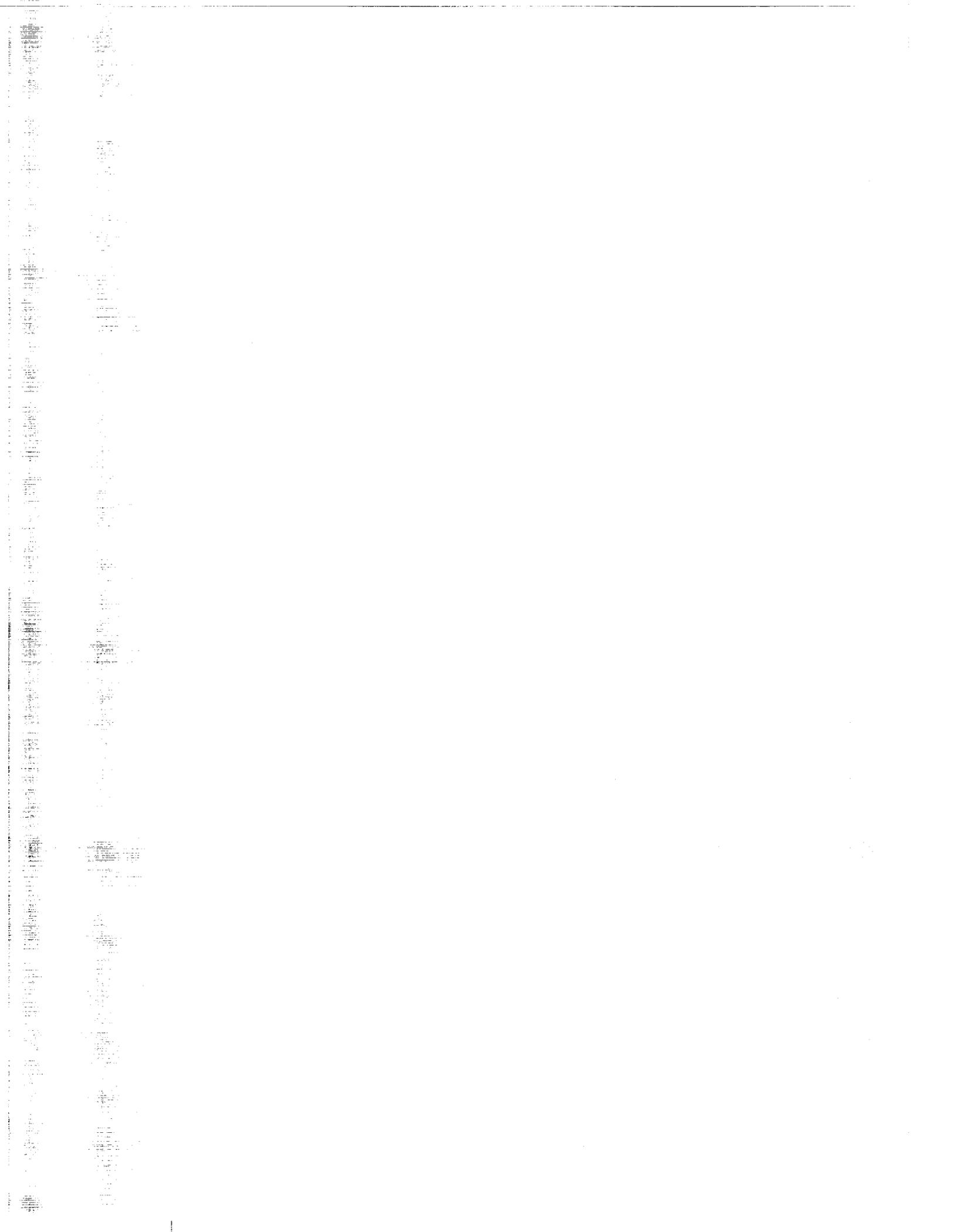
Fifth Annual Workshop on Space Operations Applications and Research (SOAR '91)

Kumar Krishen, *Editor*
NASA Lyndon B. Johnson Space Center
Houston, Texas

I

INTRODUCTION

Kumar Krishen, Ph.D.

Operations[1], as the term applies to the Nation's civil defense and space programs, constitutes a broad spectrum of activities and associated facilities that enable the conduct of a program or a mission to achieve the desired goals or objectives. They include Earth-based, in-flight, in-space, and planetary surface-based operations. Mission plans, schedules, and logistics are integral parts of operations. With regard to the conduct of the Space Shuttle Program, NASA has become increasingly conscious of the need for improving operations efficiency. The goal of operations efficiency efforts is to provide systems, services, and the infrastructure to enable safe operations at a substantially reduced cost. The operations Work Breakdown Structure typically includes space automation and robotics, training systems, in-space operations, ground operations, and associated information and communications infrastructure. In particular, recent advances in software technology and knowledge engineering are deemed crucial in providing revolutionary capabilities for operations associated with the space programs.

The National Aeronautics and Space Administration (NASA) and the U.S. Air Force have recognized the need for continued interaction in the area of space operations technology and formed the Space Operations Technology Subcommittee (SOTS) under the NASA/Air Force Space Technology Interdependency Group (STIG). The membership of the SOTS is listed in Table I.

The goals of SOTS include the following:

- Interchange technical and programmatic information related to space operations
- Share lessons learned
- Identify areas of common/mutual interest
- Encourage interdependent programs

In the past 6 years, the SOTS has identified research and applications areas with significant potential for Air Force and NASA programs. These areas are intelligent systems, automation and robotics, life sciences, environmental interactions, and human factors.

The focus of the SOTS is on the research and technology areas, which have applications to both NASA and the Air Force. Thus, the coordination and joint pursuit in these areas of overlap (see figure 1) would provide savings for both agencies.

In addition to program reviews, meetings, and written communication, the SOTS conducts the Space Operations, Applications and Research (SOAR) symposium and exhibition annually. This symposium and exhibition has become an invaluable tool to review the progress made in existing joint programs and to identify new areas for joint or collaborative efforts. Table II presents the program overview and personnel responsible for SOAR '91.

This document contains papers presented at the Space Operations, Applications and Research Symposium, hosted by the NASA Johnson Space Center (JSC) and held at JSC in Houston, Texas, on July 9 - 11, 1991. More than 110 papers were presented at the Symposium, sponsored by the U.S. Air Force Phillips Laboratory, the University of Houston-Clear Lake, and NASA JSC. The technical

[1]"Advanced Technologies for NASA Space Programs," K. Krishen, Proceedings of the 10th Annual International Space Development Conference, San Antonio, Texas, May 17-22, 1991.

PRECEDING PAGE BLANK NOT FILMED

areas covered were Intelligent Systems, Automation and Robotics, Human Factors and Life Sciences, and Environmental Interactions. The U.S. Air Force and NASA programmatic overviews and panel discussions were also held in each technical area. A keynote session chaired by Dr. Aaron Cohen, Director of the Johnson Space Center, was organized to provide Agencywide perspective of technology programs for both the Air Force and NASA. The keynote addresses of Maj. Gen. Robert Rankine, Jr., and Mr. Arnold Aldrich, who serve as the co-chairmen for the STIG, are included in this document. These proceedings, along with the comments and suggestions made by the panelists and keynote speakers, will be used in assessing the progress made in joint USAF/NASA projects and activities. Furthermore, future collaborative/joint programs will also be identified. The SOAR '91 Symposium and Exhibition is the responsibility of the Space Operations Technology Subcommittee (SOTS) of the USAF/NASA Space Technology Interdependency Group (STIG). The Symposium proceedings include papers covering various disciplines presented by experts from NASA, the Air Force, universities and industry.

TABLE I. MEMBERSHIP OF SOTS

Phillips Laboratory	**Marshall Space Flight Center**
Melvin Rogers, Co-chairman	Mr. E.C. Smith
Capt. Jim Skinner	
Lt. Col. Gale Nelson	**Kennedy Space Center**
	Mr. Tom Davis
Armstrong Laboratory	
Col. Donald Spoon	**Johnson Space Center**
Capt. Ron Julian	Dr. Kumar Krishen, Co-chairman
Dr. Samuel G. Schiflett	Mr. Robert Savely
	Dr. Howard Schneider
SDIO/ES	
Mr. Richard Iliff	**Jet Propulsion Laboratory**
	Mr. Wayne Schober
AFOSR/NM	
Dr. Abe Waksman	**Langley Research Center**
	Mr. Jack Pennington
Wright Laboratory	
Capt. Mike Wellman	**Ames Research Center**
	Mr. Allen Fernquist
AFOSRL/NL	Dr. Michael Shafto
Dr. John Tangney	Dr. Mary Connors
NASA Headquarters	**Lewis Research Center**
Mr. Mark Gersh	Dr. Dale Ferguson
Mr. Geoff Giffin	

X101325M

TABLE II. PROGRAM AND RESPONSIBLE PERSONNEL

Program

SOAR '91 will include USAF and NASA programmatic overviews, panel sessions, exhibits, and technical papers in the following areas:

- Intelligent Systems
- Automation and Robotics
- Life Sciences
- Human Factors
- Environmental Interactions

Technical Sessions and Exhibit Hours

Tuesday, July 9	8:00 am - 7:00 pm
Wednesday, July 10	8:00 am - 7:00 pm
Thursday, July 11	8:00 am - Noon

Welcome/Opening Addresses (July 9, 8:30 - 9:30)

NASA/Air Force
Mr. Geoff Giffin - NASA Headquarters
Dr. Allan Schell - Air Force Systems Command

Panel Discussion (July 9, 3:30 - 5:00)
Technology Requirements

Moderator: Dr. Kumar Krishen - NASA/JSC
Panelists: Geoff Giffin, Mr. Peter Ahlf - NASA/HQ,
Mr. James Romero, Col. Ray Barker/USAF

Keynote Session (July 10, 6:30 - 9:30)

Master of Ceremonies: Dr. Aaron Cohen - NASA/JSC
Keynote Speakers: Major General Robert Rankine, Jr. -
Air Force Systems Command and Mr. Arnold Aldrich -
NASA/HQ

Exhibitors

Computer Sciences Corp.	McDonnell Douglas
Deneb Robotics	NASA/Ames Research Center
Digital Equipment	NASA/Johnson Space Center
EXOS	Oracle Corp.
NASA/Goddard Space	Rice University
Flight Center	Rockwell-Downey
Grumman Corp.	Silicon Graphics
Hewlett Packard	Space Industries
IBM Federal Sector	Template Graphics
IntelliCorp	Togai InfraLogic
KRUG Life Sciences	University of Lowell
Lockheed	USAF
LinCom Corp.	

Symposium Coordinators

Symposium General Chair:	• Dr. Kumar Krishen, NASA/JSC
Assistant General Chair:	• Mr. Mel Rogers, Phillips Laboratory
Technical Coordinators:	• Capt. Jim Skinner, Phillips Laboratory
	• Mr. Robert T. Savely, NASA/JSC
Administrative Co-Chairs:	• Ms. Carla Armstrong, Barrios Technology, Inc.
	• Dr. Glenn Freedman, University of Houston - Clear Lake
Exhibit Co-Chairs:	• Mr. Charles Pittman, NASA/JSC
	• Mr. Ellis Henry, I-NET, Inc.
	• Ms. Bette Benson, University of Houston - Clear Lake

Technical Area Coordinators

	NASA	USAF
Intelligent Systems	Mr. Mark Gersh NASA/HQ	Capt. Jim Skinner Phillips Laboratory
Automation and Robotics	Mr. Jack Pennington NASA/LaRC	Capt. Ron Julian Armstrong Laboratory
Life Sciences	Dr. Howard Schneider NSA/JSC	Dr. Samuel G. Schiflett Armstrong Laboratory
Human Factors	Ms. Mary Connors NASA/ARC	Col. Donald Spoon Armstrong Laboratory
Environmental Interactions	Dr. Dale Ferguson NASA/LeRC	Lt. Col. Gale Nelson Phillips Laboratory

Operations Technologies and Research

NASA Operations, Research and Technology

Air Force Operations, Research and Technology

SOTS Focus

Space operations technology applications and research of mutual interest to the Air Force and NASA

X101324M

Figure 1. The SOTS Domain of Technologies

MESSAGES

General Chair,
Dr. Kumar Krishen
NASA Johnson Space Center

The goals of the Space Operations Technology Subcommittee (SOTS) of the Space Technology Interdependency Group (STIG) include interchange of technical and programmatic information, sharing of lessons learned, and the identification of interdependent programs related to space operations. The SOAR Symposium and Exposition has been, and continues to be, an excellent means to accomplish most of the SOTS goals. This is the fifth time the Air Force and NASA will host the Symposium and Exposition to evaluate progress of our ongoing efforts and identify future cooperative programs. To this end, exhibits, technical papers, panel discussions, and programmatic reviews have been planned for the technical interchange. The present fiscal climate of our nation makes it incumbent upon us to avoid duplication and embark on cooperative and joint technology development for Air Force and NASA applications. With your participation, I believe we can achieve this coveted goal and advance the future programs of both NASA and the Air Force.

Assistant General Chair,
Mel Rogers,
Kirtland AFB

As the new Air Force Co-Chairman for the Space Operations Technology Subcommittee, I would like to welcome your participation in the 1991 SOAR Conference. The past year has been a dynamic one within the Air Force. We have seen the restructuring of the Air Force. We have seen the restructuring of the Air Force laboratories and the implementation of Project Reliance to increase cooperation between the services. As a result of the restructuring, the Phillips Laboratory has been designated as the new Air Force superlab dedicated to developing technologies for space. Our participation in past SOAR conferences has enabled Air Force and NASA program and project managers to reduce the duplication of effort and revealed new application areas for existing technologies. In addition, it has given the experts from both organizations the opportunity to exchange information about the technologies that are critical to space operations. The contacts that have resulted from these interchanges are a valuable resource that lasts year 'round. This year's workshop promises to be another excellent forum to exchange technical information and identify opportunities for joint and cooperative ventures related to space. I look forward to meeting with all of you at SOAR '91. I know that with your participation the conference will be informative and productive.

Maj. General Robert R. Rankine

MAINTAINING TECHNOLOGICAL AND INDUSTRIAL SUPERIORITY

Ladies and Gentlemen, it is an honor for me to be with you this evening. I would like to begin my opening remarks by sharing with you a reminder from our late president—John F. Kennedy:

> "The most powerful single force in the world today is neither communism nor capitalism; neither the H-Bomb nor the guided missile—it is man's eternal desire to be free."

Freedom is the greatest gift a nation can offer its people. In a world plagued by constant turmoil and by war, in a world where most of the population is ruled by dictatorship, America's technological and industrial superiority has sustained the beacon of freedom throughout the world, providing all of us hope for a better future.

America won the cold war with technology. We won decisively in the Kuwaiti theater. However, the overall U.S. lead in technology relative to the rest of the world has eroded over the last 20 years, and the outlook is for greater technological competition in the future. In the years ahead, the U.S. will face major challenges resulting from increased foreign competition and erosion of our technological and industrial base.

Our declining industrial competitiveness is a potential source of international instability, a threat to our industrial self-sufficiency, both military and commercial, and the key to becoming a second-rate economy. The prospects of increased international competition and the relative decline in our technological and industrial base have recently received high-level attention. For example, according to Donald Atwood, Deputy Secretary of Defense, "The deterioration of America's industrial base is one of the most pressing issues facing the Department of Defense."

To prevent the continual erosion of our industrial base, we must enhance our technology base. As funds decline, technological superiority will become the backbone of deterrence and stability in the 1990s. Foreign competition, the relative decline in our industrial base, and the combined problems of technology diffusion and weapons proliferation will all make it that much harder to keep our technical edge. We must plan carefully, finding ways to do our business better, and advance technologies that will ensure America's future security and economic growth. Space is one of the major technological areas that will accomplish this goal and provide us with the greatest future flexibility.

The importance of space for national defense, technological, and industrial growth has grown—and will continue to grow. If this country is to maintain its historic leadership role, our commitment and dedication to space activities must continue to expand. In fact, as we have seen from "Operation Just Cause" and from "Desert Storm," space systems have become an integral part of military operations. Rather prophetically, several months prior to the incidents that led up to Desert Storm, Secretary Rice remarked:

> "Space is a natural extension of the Air Force's operating medium. In an unstable world with refocused threats, space offers stability and control...to be a 21st century superpower, the United States needs the ability to help friends and quell enemies within hours. Only with aerospace forces can you concentrate and reconcentrate power that quickly."

These words underscore the power and value of space technology in securing a strong and safe future for America. But space is not just a military resource, it is an invaluable resource to commerce and it is critical to our economic future. America's projected investment in space for 1991 is approximately $30 billion. NASA and the Department of Defense will spend nearly $27 billion, and industry will spend almost $3.6 billion purchasing and launching satellites. The utilization of space as a resource is realistically still in its infancy, yet the benefits derived from this resource profoundly affect the world around us.

Space technology and development will become as critical to our country's future as the industrial revolution was critical to our country's growth in the 18th century. Future growth of American industry, technology, and political strength are inextricably tied to space. America is tied to space because of its importance to both the national security and civil space sectors—it helps maintain the peace and enhances the life of every American living today. From communications to entertainment, surveillance to navigation, weather forecasting to storm warnings, science to space station, and the list goes on and on, space is central to American development today and tomorrow.

The problem that confronts all of us is "How do we sustain progress in defense technology—and, in particular, in defense space technology—in a time of declining defense budgets?" One way is to streamline in order to get more out of our limited funds. To achieve this goal, Mr. Richard Cheney, our Secretary of Defense, instituted the Defense Management Review process. Subsequently, Deputy Secretary of Defense Atwood challenged the Military Services by issuing a draft Defense Management Report Decision. This draft report decision, DMRD 922, focused on the Consolidation of Research and Development Laboratories and Test and Engineering Facilities of the military services. DMRD 922 became the catalyst that sparked formal discussions on ways to enhance cooperation and increase reliance among Service programs and facilities for research and testing.

Responding to the challenges of the DMRD 922, the Services each initiated its own laboratory consolidation proposal and collectively, the Services initiated Project Reliance to place more dependence on one another's capabilities. In response to the DMRD 922, the Air Force has taken two major steps to streamline its laboratory management structure and improve the quality and relevance of Air Force science and technology programs. The first step consolidated our scientific and technology resources at 14 different laboratories and research centers into four major laboratories. This realignment facilitates greater interaction between scientists and engineers who must develop the technology for the interdisciplinary Air Force weapon systems of the future. Our "super" laboratories will also reduce management overhead. At the same time, simpler laboratory organization will clarify and focus our missions and create more opportunities for Tri-Service Reliance and external recognition.

The four major laboratories were organized around the applied technologies that serve their parent product division. Armstrong Laboratory at Brooks AFB in Texas supports "human systems; Rome Laboratory at Griffiss AFB in New York supports "command, control, communications, and intelligence"; Wright Laboratory at Wright-Patterson AFB in Ohio supports "air vehicles and armament"; and the Phillips Laboratory at Kirtland AFB in New Mexico concentrates on space, missiles, and directed energy technology. Each of the four "super" laboratories have also been assigned certain corporate research responsibilities, such as materials research at Wright Laboratory and geophysics research at Phillips Laboratory. This reorganization will enhance our technology, help us meet our user's needs, and improve our acquisition capabilities through the increased synergy and teamwork that will bring our people, processes, laboratories, and their SPO customers closer together. As an end product, we look forward to greater combat capability for the Air Force and the Nation.

The second major step we've taken established an investment planning process to ensure the quality and relevance of our Science and Technology programs. Our Air Force acquisition executive, Assistant Secretary Jack Welch, requested the Air Force Scientific Advisory Board to review all Air Force Science and Technology programs, and we have further asked the Board to explicitly evaluate

and score those programs based on their technical quality. Similarly, we have asked the Air Force's operational commands to score the importance of our major technology demonstration programs. Using this feedback from the Air Force Scientific Advisory Board for technical quality and the Air Force using commands to determine the operational relevance, we will keep a balanced focus on customer needs while advancing technology innovation in new research areas.

Our vision for Air Force laboratories will ensure they remain world class organizations working in technical areas vitally important to the Nation's economy and national defense. But we must also enhance cooperation and increase reliance among all government research organizations if we are to maximize our available resources to strengthen our national technological capabilities.

During the summer of 1990, the three Services presented a coordinated proposal to Mr. Atwood that outlined several approaches to laboratory consolidation and inter-Service Reliance. Once Mr. Atwood conceptually approved of the Tri-Service proposal, the Army, Navy, and Air Force implemented the Tri-Service Project Reliance process to identify ways of achieving inter-Service science and technology consolidation.

By November, Mr. Atwood signed the final version of DMRD 922 and directed the Acquisition Undersecretary to provide a plan outlining management actions to implement DMRD 922. Subsequently, in December 1990, the Joint Logistics Commanders institutionalized the Project Reliance philosophy and process through the Joint Directors of Laboratories. The charter of the Joint Directors of Laboratories has been broadened in scope to conduct inter-Service Reliance analysis and joint technology planning on a continuing basis and to implement the results.

Before Project Reliance, the primary method of interaction among the Services was coordination and exchange of information at the working level. The Tri-Service Reliance activity is a comprehensive effort to identify means for increasing inter-Service Reliance in Science and Technology programs and to give the resulting joint programs visibility at the Director of Defense Research and Engineering level. This effort has yielded substantial positive changes for achieving increased efficiencies through collocations, consolidations, and joint Service S&T programs. Formalized Service agreements have been established for each of the technologies considered that were not Service unique. As a result of these agreements,

- There are 71 technology sub-areas where the Services will jointly plan for work to be conducted at separate Service locations.

- There are 105 technology sub-areas where work will be collocated to various single-Service sites for program execution.

- There are 4 major technology areas and 6 technology sub-areas where programs will be consolidated and carried out at a single location under a lead service for management.

The four objectives of continuing oversight of Tri-Service science and technology by the Joint Directors of Laboratories are:

1. Structuring an integrated DoD S&T program in selected technology areas through joint program planning and execution;

2. Establishing technology panels to plan and oversee execution of cooperative Service S&T programs;

3. Collocating S&T efforts to achieve critical mass and avoid unwarranted overlap and eliminate duplication among Service programs; and

4. Conducting annual joint and corporate Science and Technology reviews for the Director of Defense Research and Engineering and his staff

These activities will enable all Services to increase their efficiencies in areas of common interest without severing the close ties to the end users of the technology and will free up resources, which each Service can reinvest in critical, higher-priority technology areas. These changes will reduce the perception of program fragmentation and duplication, promote an integrated "joint" Tri-Service program, and enhance the overall quality of DoD's S&T program. through this improved strategic planning process, we will have greater control over our technological future.

We now wish to extend this collaborative spirit to other Federal agencies that invest in science and technology that is relevant to defense needs. By expanding the membership of the previously established NASA/Air Force space Technology Interdependency Group to include the Army and the Navy, we have created a forum for Tri-Service cooperation with NASA in space technology. That is why we are currently holding this combined Space Technology Interdependency Group and Joint Directors of Laboratories meeting, to couple previously successful Tri-Service cooperation with previously successful Air Force/NASA cooperation.

In the past year, the Services have made remarkable progress in establishing a new era of Science and Technology cooperation. IntraService laboratory consolidation reduced our management overhead, helped us focus our efforts, and increased our Tri-Service Reliance. The changes within the Army, Navy, and Air Force Laboratories have greatly strengthened our technical capabilities. Tri-Service Science and Technology cooperation has never been better. We wish to expand this Science and Technology cooperation to other federal agencies beginning with NASA. Working together as a team, we should be capable of maintaining our world leadership role in space technology. In fact, by working as a team, we should be able to far exceed the objective of merely maintaining a world leadership role; we should be able to lead our nation into a new golden era of space technology and applications.

Arnold D. Aldrich
NASA Associate Administrator
Office of Aeronautics
Exploration and Technology

COLUMBUS, VISIONARIES, AND BOLD ACTION

The past is beyond our ability to change, but the future is ours to make, as we see fit. Christopher Columbus, in his time, was in no way able to visualize the future on a five-hundred year time horizon, and we who will soon celebrate his far-reaching achievement are not much better prepared to visualize our future on such a long-range time scale.

Columbus and his small band of individuals did, however, in their courageous pursuit of a vision of new routes to the Orient, instead open up new avenues to economic prosperity and thrust the world forward into the current era. With the daring to press beyond the constraints of their time, they activated a process that led to acquiring new levels of knowledge about the world in which we live.

That same spirit was alive and well a hundred years later when the colonization of North America began, and it was even stronger at the end of the 18th Century when the United States was formed and the task of building civilization across this continent moved forward in earnest. Although it has been many years since the western frontier vanished, that heritage, the heritage of the explorer and the builder, endures within America today.

Nowhere today does that spirit burn more brightly, nor is the legacy of the explorer and the builder embodied more profoundly than in the United States space program. And it is within this space program, with the men and women of NASA, the Department of Defense, the Department of Energy, the aerospace industry, and the academic research community, that we, as a nation, have the opportunity to adopt a vision and initiate bold action for the future. If the United States is to prosper in the next century, we must have a guiding perception of the future as it could be, and work today to build towards those goals.

For the present, our nation's space program is leading the world, but in the next century, with the pace of current international competition, it could well be relegated to second class, or even to a lesser status. George Santayanna said that those who fail to learn from the mistakes of history are doomed to repeat them. Tragically, history shows that those societies that fail to aspire to a challenging vision of the future will realize a self-fulfilling prophecy of diminishing accomplishments.

Twenty-five years from now, in the year 2016, we could be achieving exciting, inspiring successes in space. The frontiers of space science, the return to the Moon, the exploration of Mars, and understanding Earth itself all challenge us. But what, in fact, will we actually accomplish over this timeframe?

I would like to remind you of the motion picture "2001 - A Space Odyssey." Created in the 1960s, Arthur C. Clarke's vision of a possible future in space stands largely unfulfilled. That vision was shared by towering figures of the early U.S. space program such as Webb, Gilruth, Von Braun, and the other leaders who guided the Agency through the bold space adventures of the 1960s, 1970s and the early 1980s.

The question we must ask ourselves now, 30 years later, is whether we still have the vision of a nation of explorers and builders? If we envision great accomplishments in the coming century, what steps must we take in this decade to make those goals achievable? Conceiving or embracing inspired visions is not enough; we must also act with conviction.

I believe a central, crucial path of action is to build on our outstanding capabilities in technological research and innovation. We must make a renewed commitment to the development of advanced space technologies. To expand our frontiers in space, we must challenge established technologies and explore new approaches to our space endeavors. We must find better ways if we are to achieve greater heights.

It has been widely demonstrated that the space program is both *driven by* and also is a proven *driver of* technology. Just as the explorations of past centuries have resulted in tremendous economic benefits, so too can our space program of the next century provide equally vital returns.

However, we must also have the courage to acknowledge our present difficulties, and we must have the determination to move effectively to overcome them. Despite our past accomplishments, today the art of space engineering is relatively new and immature, comparable perhaps to the state of engineering in aeronautics in the 1930s or perhaps even the 1920s. As a consequence, we have the need for a broad range of technological advancements that must be vigorously pursued.

Many national leaders have articulated their concerns over the current status and outlook for the United States space program and have emphasized the need for stronger investments in research and technology to enhance our competitiveness. Most recently, the report by the Advisory Committee on the Future of the U.S. Space Program, otherwise known as the "Augustine Committee," stated that NASA should provide a stable share of its budget for space technology and fund it at two or three times the current level of investment. They also recommended that NASA formulate an agencywide technology plan and subject it to a broad review by experts external to the Agency.

We are currently deeply immersed in formulating a comprehensive "Integrated Technology Plan" (ITP) which is responsive to these recommendations. And we have just completed a week-long review of this plan by specialists from industry, academia, and other agencies. However, we recognize that in the world of budget realities, it may be some years before NASA can fully support an expanded space technology program at the level recommended by the Augustine Report.

Critical, technology challenges are highlighted in the ITP, technologies that have been emphasized by the Augustine and Synthesis reviews, as well as by our own 90-Day Study. Some of the major challenges are the development of technologies for

- Space nuclear power and propulsion,

- Heavy lift launchers and orbital transfer vehicles,

- Information collection, processing, and visualization for data quantities equivalent to multiples of the Library of Congress, and

- A myriad of human support systems, including regenerative life support and EVA systems for both in-space and surface operations.

As we are all well aware, space technology budgets are under pressure, not only in the NASA civil space program but also in the Department of Defense, the Department of Energy, and in other agencies as well. Shortfalls in these areas, in turn, adversely affect industry's Independent Research and Development (IR&D) budgets and university research grants and contracts. Thus, this becomes, more than ever, a time to leverage all national technological resources by communicating and,

wherever possible, coordinating research and technology program goals, plans, and products. It is a time for partnerships in the national interest within government and between government, industrial, and academic sectors.

The list of space program challenges and opportunities is formidable, and many of them can be effectively addressed through advances in technology. In particular, the costs of year-in and year-out space operations must be actively pursued. Launch costs are too high, and launch operations are complex and time-consuming. If we are to achieve our goals in science and in exploration, we must have both low cost, reliable, cargo transportation to orbit and assured access to space for human operations. Together, steps to control operations costs and adequate near-term investments in advanced technology and transportation systems can provide the foundation for the space program of the future.

The DoD vision of achieving military technological superiority, so aptly demonstrated in the bold U.S. response in Desert Storm, is symbolic of what this nation is capable of accomplishing. Today, the civil space program must have the vision to see beyond the constraints of our time, to envision how we can open new routes to national success and economic prosperity and to help move this nation to the forefront of a new era. There's so much out there that needs doing, so many science and exploration goals yet to be achieved, so much that is still unknown. That is what is exciting. We must anticipate what the future could be, and more importantly, we must resolutely take the bold steps in technology necessary to ensure that that future comes to pass.

The occasion of this banquet is certainly a tribute to the potential for cooperation in space technology activities between many contributing sectors of our industry. Tomorrow and Friday the Space Technology Interdependency Group, STIG, takes a visionary step toward becoming a broader federal forum for technology cooperation with the addition of U.S. Army and U.S. Navy membership, and there will also be informal participation by representatives from DARPA, SDIO, and DOE.

As you have heard from General Rankine, the tri-services technology panels, under the Joint Directors of Laboratories, have met here today with a focus on space activity. And the week's underpinning event is the Space Operations, Applications, and Research (SOAR) conference. This activity, conceived by the STIG's Space Operations Committee has, for the past five years, created a cooperative forum for NASA and the Air Force to reach out to industry and academia.

I congratulate all of you, the leadership and participants in this "triple-header" demonstration of national technological cohesiveness, and I believe these activities will greatly assist in building the cooperative relationships so essential to the realization of each of our organizational visions. These activities serve as a model for the leveraging of total U.S. capabilities toward the goal of continuing space program preeminence.

CONTENTS

PRECEDING PAGE BLANK NOT FILMED

Session I4:   **ICAT**
Session Chair: **Jim Fleming**

Session I5:   **MISSION OPERATIONS**
Session Chair: **Lynne Cooper**

Session I6:   **SOFTWARE ENGINEERING**
Session Chair: **Dr. Kirstie Bellman**

Session I7:   **PLANNING AND SCHEDULING**
Session Chair: **Chris Culbert**

SECTION II: AUTOMATION AND ROBOTICS

Session A1:   AUTOMATION AND ROBOTICS PROGRAMMATIC SESSION
Session Chair:

Session A2:   SENSING AND DISPLAY
Session Chair: Prof. Neil Duffie

Session A3:   ERROR DETECTION AND CONTROL
Session Chair: Joe Herndon

Session A4:   MANNED PERFORMANCE
Session Chair: Dr. Mark Stuart

Session A5:   **TELEROBOTICS AND MECHANISMS**
Session Chair: **Prof. Mark Friedman**

Session A6:   **SYSTEM PLANNING**
Session Chair: **Jack Pennington**

Session A7:   **ASSEMBLY AND SERVICING**
Session Chair: **Dr. Charles Woolley**

SECTION III: HUMAN FACTORS AND LIFE SCIENCES

Session H5: **VIRTUAL REALITY**
Session Chair: **Dr. Steve Ellis**

Session H6: **SPACE PERCEPTION AND PHYSIOLOGY - II**
Session Chair: **Dr. Barbara Stegmann**

Session H7: **SPACE CABIN CONTAMINANTS**
Session Chair: **Dr. Jeffrey W. Fisher**

SECTION IV: ENVIRONMENTAL INTERACTIONS

|

SECTION III

HUMAN FACTORS AND LIFE SCIENCES

Session H1: MEASUREMENTS, TOOLS, AND ANALYSIS - I

Session Chair: Dr. Mary Connors

NASA Human Factors Programmatic Overview

Mary M. Connors

NASA Ames Research Center

Moffett Field, CA 94035-1000

INTRODUCTION

Human factors addresses humans in their active and interactive capacities, i.e., in the mental and physical activities they perform and in the contributions they make to achieving the goals of the mission. The overall goal of space human factors in NASA is to support the safety, productivity and reliability of both the on-board crew and the ground support staff. Safety and reliability are fundamental requirements that human factors shares with other disciplines, while productivity represents the defining contribution of the human factors discipline.

PROGRAMMATIC ORGANIZATION AND DIRECTION

Space Human Factors in NASA forms an essential component of two NASA HQ program offices, each with a complementary role to play. The Life Science Division of the Office of Space Science and Applications (OSSA) looks at human factors from a scientific perspective including connections with the biomedical sciences. The Life Science Division develops the human requirements that must be addressed in space-flight, including those associated with behavior and performance. The Information Science and Human Factors Division of the Office of Aeronautics, Exploration and Technology (OAET) is concerned with human factors as it relates to the design, development and implementation of technology products. Here, an understanding of humans' ability to use and benefit from facilities and tools is a necessary element of the technology research and development process.

Human Factors has a special relationship to the mission activities that must be performed, and Human Factors programs begin with discipline-oriented mission analyses. The general content of the OSSA program is given in Figure 1. This figure indicates the emphasis on basic understanding of human requirements and capabilities as well as countermeasures to deal with human limitations or negative effects due to the space environment. Similar information for the OAET program is given in Figure 2, indicating an emphasis on crew station design and on human performance modeling, instrumentation for crew support, and integrated systems.

Life Sciences

Space
Human Factors

Psychophysiological Interactions
Fatigue/Sleep/CR
Behavioral Processes
Performance Capability
Workload Measurement
Habitability
Crew Factors
Human/Machine Interaction
Selection/Training
Stress/Crew Support
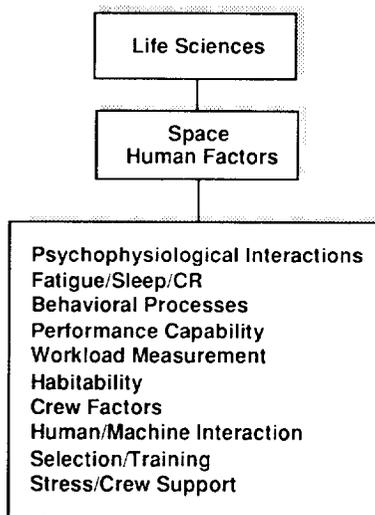
Figure 1 - OSSA Human Factors Emphases

Figure 2 also shows the OAET thrusts to which these human factors elements relate. Since human factors is a cross-cutting technology, the program impacts all thrust areas. However, human factors has a special relationship to Operations (where the actions of the human participants are central) and to Planetary Surface Explorations (where, because of distance, duration, and multiplicity of environments, the demands on human participants take on special importance).
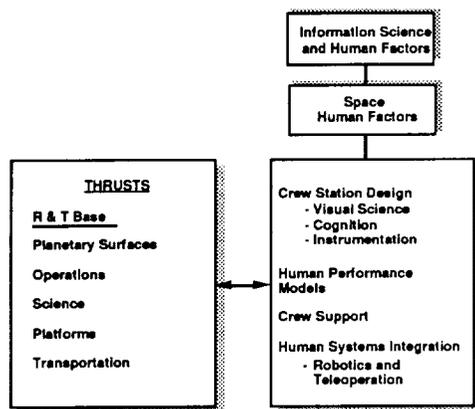
Figure 2 - OAET Human Factors Emphases

FIELD CENTER EMPHASES

Research within NASA is conducted primarily by and through the field centers. Much of this research is conducted in-house, with a significant amount being conducted through university or other granting arrangements. Human factors activity can be found at all of the major NASA centers. However, most space human factors research and development is concentrated between two centers: Ames Research Center (ARC), Moffett Field, CA, and Johnson Space Center (JSC), Houston TX. Work at the two centers complements each other, both in the nature of the approach taken and in areas of special emphasis and expertise. At ARC, the tendency is towards more development-based research, allowing the building blocks of understanding to be established, the extent and limitations of results to be assessed, and the space application to be developed. At JSC, the tendency is towards more implementation-based research, making a close connection with particular flight needs, and utilizing the Center's access to space simulation facilities and to the astronaut corps. Neither center presents a pure model but reflects, rather, a tendency or general approach. Considerations that drive the emphases also influence the location of specialty areas within the human factors discipline.

In addition to its own internal planning, NASA regularly seeks the advice of special panels to help the Agency identify important issues and needs. These panels are comprised of university, industry, and government specialists. One such panel is the Space Science and Technology Advisory Committee (SSTAC). Figure 3 presents an abbreviated description of areas recently identified by the Human Factors subcommittee of SSTAC as representing important and timely research undertakings. All of the areas identified by SSTAC are contained within the broad structure of the combined Agency space human factors program, and their annunciation by this subcommittee serves both to confirm the approach being taken and to give greater specificity to the identified research needs and opportunities.

Figure 3 - High Priority Issues as Defined by SSTAC

Figure 4 - Field Center Emphases

Human factors specialties for both ARC and JSC are given in Figure 4. At ARC, space research on crew fatigue and circadian rhythm as well as crew factors (crew communication and coordination) are natural outgrowths of ongoing aviation crew research, while in-house expertise in perceptual and cognitive sciences provide the

435

foundation for development of crew support tools such as virtual environments and 3-D audio displays. At JSC, Center responsibility for selection, training, and crew well-being translate into research to support these requirements and to the establishment of standards to guarantee them. Both centers are strongly involved in developing models to understand and design to human requirements, while both also emphasize the emerging and dominating importance of human/computer interactions. Laboratory facilities at each center, reflecting the research areas described, are listed in Figure 5.

ARC Labs	JSC Labs
Visual Science and Technology	Antropometry and Biomechanics
Cognition	Human/Computer Interaction
Virtual Reality	Graphics Analysis
Crew Factors	Remote Operators Interaction
Circadian Factors and Countermeasures	Weightlessness Environment Test Facility
Aviation Simulators/Analogs	Shuttle Simulation and Training Facility

Figure 5 - Facilities at ARC and JSC

FROM RESEARCH TO OPERATIONS

A ubiquitous problem for all research organizations involves transitioning research results for use by real people doing real jobs. This is a highly complex and dynamic undertaking that requires, at a minimum, establishing a partnership between two quite diverse communities. The importance and difficulties of moving from research to operations has already been raised several times in this conference and deserves considerably more attention. However, I would like to concentrate here, not on the broad transition issue but on the narrower and slightly different issue of how researchers themselves might think about and structure their activities to better position themselves (and their findings) for eventual application to space.

Human factors researchers, like other researchers, select their methods and approaches based on the demands of the questions to be answered. Some questions are best addressed in the highly controlled environment of the laboratory; others address contextual factors and must be studied in the field. Frequently, a variety of approaches is needed to fully understand the contributions of the various elements. Figure 6 lists a number of research approaches.
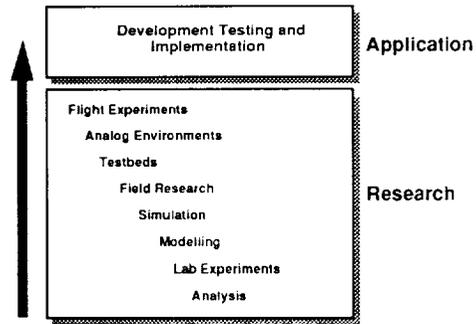
Figure 6 - Research Approaches

For the space human factors researcher the task is to plan his or her activities to "move" the research to where it can inform spaceflight. This means both understanding the aspect of human behavior being studied and understanding how environments might change this behavior. The former without the latter is unusable; the latter without the former is unsupportable. Taking one example, a phenomenon might be examined initially in the laboratory, modeled and verified for generalizability, and tested in the field. At this point, a controlled investigation in an analog environment or in space may be indicated. The specific progression of the research will change with the subject of investigation. However, in all cases the mental set of the researcher must include both building a solid scientific understanding and progressing towards the space application.

SOAR '91

It is not possible in any one conference to discuss all aspects of the NASA Space Human Factors program as outlined in general terms above. In this year's SOAR Conference, a small number of topics have been selected for inclusion in the Human Factors/Life Science track. The topics selected fit the NASA perspective in the following ways. The two sessions on Methods, Tools, and Analysis provide an opportunity to discuss and highlight recent advances in structural and performance models and their utility in understanding human performance in space. The Virtual Reality session focuses our interests on this new method of individually simulating a desired environment. Such simulation can be

436

almost dimensionless and could be utilized, for instance, in support of onboard training for a Mars mission. The Human/Machine Interaction session underscores the importance of the emerging field of crew systems and the need to understand the dynamics of these systems as applied to separated and highly autonomous crews. Although these topics define only a subset of the ongoing NASA program, they are important both in their own right and reflect a confluence of interest and activity in NASA and the Air Force.

N92-22326

Development of an Empirically based Dynamic Biomechanical Strength Model.

A. Pandya(1), J. Maida (2), A. Aldridge(1), S. Hasson(3),
B. Woolford(2)

(1) Lockheed Engineering and Sciences Company, Houston, Tx.

(2) NASA-JSC, Houston, Tx.

(3) Texas Womens Univesity, Houston, Tx.

Abstract:
Computer aided engineering (CAE) is commonly used in many aspects of aerospace engineering. Extensions and enhancements of these useful tools of analysis are now beginning to be applied to the complex area of human modeling. The overall goals of such systems include analyses of the performance capabilities of a given individual or population in a specific environment. This is a multifaceted problem. The issues of anthropometric representations, kinematic articulation of joints (reach), vision, and strength are just a few examples of the areas of complexity involved. The focus of this report is on the development of a dynamic strength model for humans.

Unlike earlier attempts at strength modeling, which were based on rotational spring and damper systems, our model is based on empirical data. The shoulder, elbow, and wrist joints are characterized in terms of maximum isolated torque, position and velocity in all rotational planes. This information is reduced by a least squares regression technique into a table of single variable second degree polynomial equations determining torque as a function of position and velocity. The isolated joint torque equations are then used to compute forces resulting from a composite motion-- which, in this case, is a ratchet wrench push and pull operation. What is presented here is a comparison of the computed or predicted results of the model with the actual measured values for the composite motion.

Introduction:

Computer aided engineering (CAE) analysis tools are being applied to a wide variety of applications from light and sound raytracing, heat transfer models to finite element analysis of structures. The techniques of CAE are now beginning to be applied to the issues involved in human modeling (strength, vision, and reach analysis). What is presented here is a dynamic human strength model.

A dynamic strength model could be used to assess and predict whether a person or population is capable of performing a physical task on the job. This is important in the case of space Extra Vehicular Activities (EVA) where crew members need to handle massive structures such as satellites and various space assembly components. In these situations, mission planners would benefit from a simulation model of all the forces, torques and accelerations that would be imposed by and imposed on the crew member.

Equipment design engineers could also benefit from a strength model. Design specifications can be enhanced if engineers could predict the forces and torques to be applied on or with a given piece of equipment. These applications include, for example, threshold torques needed to open hatches and doors and to operate tools needed for assembly or to determine maximum forces applied to ensure that the equipment will not be damaged. Equipment may be better designed if information on the strength of the user population were available.

Equipment placement designs and scenarios may also be enhanced. Questions like "What is the best configuration for this body restraint relative to this tool for maximum strength?" or "Where should this hand hold be placed for the most efficient strength utilization?" could be better answered by the systematic examination of many possibilities and scenarios with the goal of defining more comfortable and safer designs.

Lastly, a strength model is useful as a tool of study to achieve a greater understanding of how the musculoskeletal system functions, of how the torques and forces are propagated, and of what the system control mechanisms and parameters are. This knowledge may lead to, for instance, better designs of robotic and manipulator systems of the future.

Our objective is to develop and validate a human dynamic strength model using empirical data.

Method:

Data collection:

The data collection effort occured over an eight week period. There were fourteen subjects, eight males and six females, ranging in ages from 21 to 28 years. Each subject was tested isokinetically for isolated upper extremity motion (shoulder, elbow, wrist) at four velocities (60, 120, 180, 240 deg/sec) and then tested with a simulated ratchet wrench maneuver at two velocities (120, 240 deg/sec).

The general procedure for evaluating all the upper extremity joint movements was the same. Torque was measured by using a Lido multi-joint testing unit (Loredan Biomedical, Inc., Davis, CA. see figure 1). The subjects

I

were positioned so that the axis of the joint was directly in-line with the axis of the dynamometer goniometer. Dynamometer attachments were selected and placed in order to isolate the joint being measured. The subject was positioned on the instrument and maximally stabilized with the joint positioned at a specified initial condition. The subject was then instructed to give a maximum effort for each of five repetitions and informed to move the isolated joint through the entire range of motion. A three minute recovery period was taken before each change in velocity setting. The axes of motion measured were the shoulder flexion-extension, shoulder medial-lateral rotation, shoulder abduction- adduction, elbow flexion-extension, wrist flexion-extension, wrist radial-ulnar deviation, wrist supination-pronation. The setups for these motions are described in the Lido multi-joint testing manuals.

For the multi-joint test, a ratchet wrench maneuver, the subject was stabilized with velcro straps at the waist and across the chest. The subject gripped a simulated ratchet device at a height of 90% of the linear distance measured from the subjects greater trochanter to the acromio-clavicular joint. The range of motion for the ratchet bar was between 45 and 50 degrees. To minimize the motion of the upper extremity, the subject extended the elbow and shoulder fully forward without bending at the waist. This test was also a maximum torque effort of five repetitions with a three minute recovery period before each change in the velocity settings (120, 240 deg/sec).

The anthropometric data which was collected included height, weight, age, sex, skinfold measures and dimensional assessment. The anthropometric data format is documented in NASA's Man Systems Integration Standards (MSIS) document (NASA-STD-3000) [6]. The standard was also used to provide the joint limit information. Joint limits for the model were applied statistically as this information was not collected in our study.

Figure 1 Loredane Inc., Lido multi joint testing system.

Data reduction:

The data was collected using the Loredan software, "Lido Active 3.3", executing on an IBM PC. For all cases, the data set consisted of torque and angle pairs. The data was uploaded to a graphics workstation (Silicon Graphics), formatted into an ASCII file, noise filtered, reformatted to aid the polynomial coefficient calculation and reduced to a table of coefficients of second degree polynomials. The polynomials coefficients were computed using a least squares regression method. The polynomial represents the torque as a function of angle. For each joint, the complete model input format consisted of the name of the joint, the axis of rotation, the direction of rotation, the number of polynomials and the list of the calculated torque polynomial coefficients for each velocity (one polynomial per velocity). The modeling program builds its internal lookup tables from this data organization (see Figure 2) .

elbow

y
extension
4

60.000000	6.281160	0.069170	0.003640
120.000000	3.587850	0.293850	-0.000580
180.000000	-1.247410	0.618840	-0.006430
240.000000	-11.244930	0.972970	-0.01040

y
flexion
4

60.000000	16.040550	0.304760	-0.000910
120.000000	15.216640	0.239190	0.000350
180.000000	16.851919	0.162090	0.000010
240.000000	8.983800	0.513420	-0.005050

Figure 2- Example input file for strength model derived from collected data for elbow for one subject.

In addition, the anthropometric data collected for each of the subjects is processed by the modeling program into a geometric human model. The specific anthropometry is necessary in order to properly convert the torques to forces for a particular individual. The human model is then made into a fully articulated human representation with proper segmentation of the body parts and statistically determined joint limits.

Environment setup:

Each individual was created in the graphics environment using that individual's anthropometric data. The initial conditions of the ratcheting operation were set to match, as closely as possible, the actual conditions. This was a critical step for validation. The main parameters of the initial conditions included the initial and final joint angles for the ratcheting motion, the distance of the hip from the rotation point of the ratchet axis and the height of the end-effector on the ratchet. Using the graphics environment, all these initial conditions were set for each individual prior to the execution of the computer simulation of the ratchet operation (figure 4).

439

Torque vector calculation.

Each joint of the upper extremity was associated with a table of polynomial coefficients describing its dynamic torque production potential [9] (figure 2). In the modeling process, the tables were loaded into computer memory for use by a table lookup module. When a joint motion occured in our test case, the axis of motion, the direction of motion, angle of motion and speed of motion were mapped to the appropriate polynomial and a torque value returned.

Since each axis of rotation for a particular joint is perpendicular to each other axis for that joint, the square root of the sum of the squares was used to determine the available torque for each joint involved in that motion.

$$Ts = sqrt(tx\ ^*tx + ty\ ^*\ ty + tz\ ^*tz)$$

where

Ts = total torque for shoulder
tx = torque for x axis
ty = torque for y axis
tz = torque for z axis

The torque values at the other joints were similarly calculated.

For each joint, the lever arm to the point of application of the force, in this case the palm, was determined. This is the Euclidean distance from the location of the center of rotation of that joint to the end-effector location. The torque values for each of the joints were converted into forces at the end-effector by dividing out the respective lever arm lengths (Ls, Le, Lw).

Fs = Ts / Ls
Fe = Te / Le
Fw = Tw / Lw

The force values were then applied to the respective direction vectors of rotation and vectorially added to produce the total end-effector force Ft. Ft represents the total force at the end-effector from the contributions of all joints in the chain. Ft is also perpendicular to the lever arm. The direction of Ft was calculated by taking the cross product of the current lever arm with the previous lever arm and then crossing the resultant with the current lever arm. This calculation was performed at each iteration for each joint.

For the test case the force vector <Ft> needed to be resolved to a torque value at the ratchet axis. This was done by first projecting the force vector <Ft> onto the normalized direction vector of rotation <Rt> for the ratchet bar yielding a vector <Fproj> in the direction of rotation of the ratchet bar with a magnitude representing the force applied in that direction. This force was then multiplied by the lever arm length (Lr) of the ratchet, the distance from the point of rotation to the point of application, yielding a torque value (Tr). This torque value and the current angle of rotation of the ratchet bar were written to a file. In addition, the force vector <Ft> and the force vector <Fproj> were graphically displayed. (see figure 3)
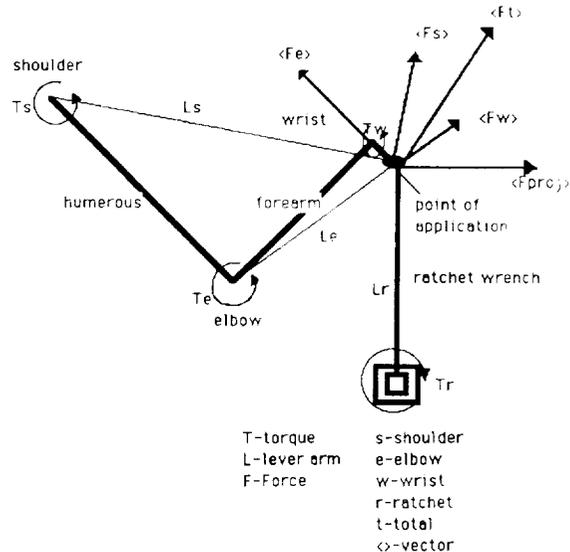
Figure 3- Diagram illustrating the force vector propagation.

Modeling the motion (Inverse kinematics):

To model the reaching characteristic of the arm while operating the ratchet bar, an inverse kinematics algorithm was needed to solve the joint angles of the arm [2,7] . Also, the human model with its corresponding anthropometry needed to be accessible to the force modeling software in order to integrate the torque functions with the motion of the arm. A software package named JACK [1], developed at the University of Pennsylvania, was used as a platform for our strength model. Although many enhancements and modifications were required, the underlying inverse kinematics and anthropomerics implementation permitted us to model the required motion.

The simulation of the ratchet bar motion consisted of the following sequence of events.

1) The parameters of the motion (start angle, end angle, steps to take, and the velocity of the ratchet) were input.

2) Time sequence information was computed which satisfied all the conditions of the ratchet's motion.

3) The location of the joint chain forming the arm and the location of the point of application on the ratchet bar were graphically selected.

4) Using the joint chain information, the torque functions for each component of the arm were loaded into the force model for use during the iterations of the ratchet operations.

5) The ratchet was moved to its initial or next position and the inverse kinematic module invoked to grasp the point of application on the ratchet with the specified end-effector (palm) in the joint chain.

6) The state of all the components of the arm, the joint angles of the arm and the state of the ratchet were extracted and input to the force model where the torque prediction was computed and written to an output file.

7) Steps 4 and 5 were repeated until five iterations of pushing and pulling were performed (See figure 4) .

440
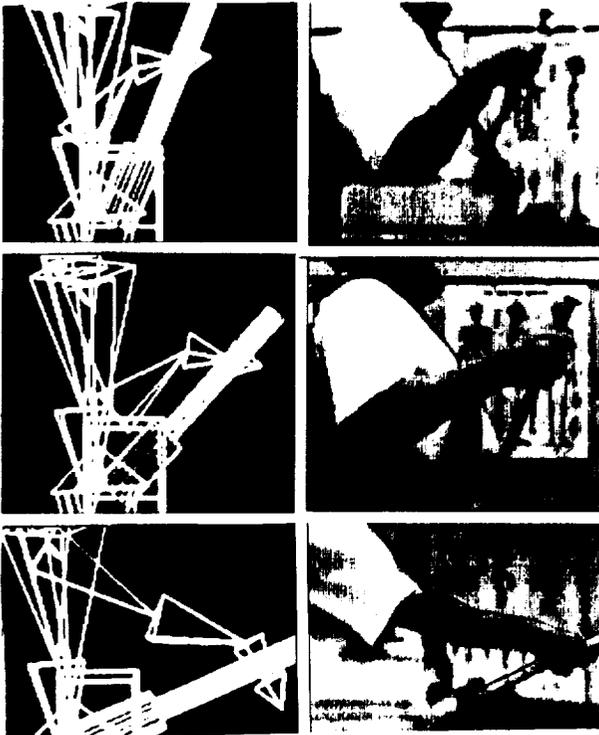
Figure 4 Comparison of actual vs. modeled ratchet wrench motion.

In order to validate the reaching motion calculated by the inverse kinematics algorithm, a real time magnetic tracking system was devised for input into the algorithm. The tracking system consisted of a Polhemus isotrack magnetic tracker connected to a Silicon Graphics Workstation. The magnetic tracker was linked to the end-effector of the man model representation. As the tracker was moved in space by a person, it fed the position and orientation information of the person's end-effector to the inverse kinematics algorithm. This information was then used to simulate the motion of the person's arm in the
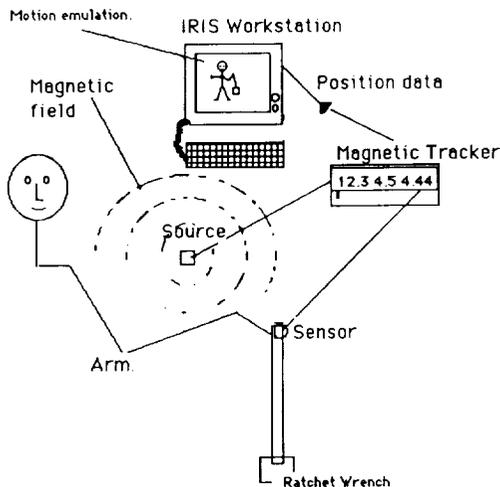
Figure 5. - Magnetic tracking setup.

computer model. The tracker of the Polhemus device was attached to a bar which could be rotated the same way the ratchet bar was rotated. Comparison by visual inspection of the actual motion of a person's arm performing a ratcheting operation with the graphically emulated motion computed by the inverse kinematics algorithm showed a strong correspondence. (see figure 5)

Analysis of data:

All the subjects were run in the graphics environment with initial conditions and orientations closely matched to the actual runs. (See Figure 4) The ratcheting was modeled at the same velocities as the measured data (120, 240 deg/sec). Output from the model were files of torque vs. angle pairs in the same range of angle values as the measured ones. For both the model output data as well as the measured ratcheting data, the average torque produced and the total work done per iteration was computed. This data was the basis of the validation of the model. Statistical analysis was done in two forms, pairwise T tests and regression analysis[8,9]. Software was written to do these tests in an automated way without user intervention.

For the T test, the measured vs. model files of the averages and total work done over all subjects were read and a difference vector is created. This difference vector is the basis of the T test comparison. Our hypothesis is that there is no difference between the means and the work between the model and the measured values. That is, assume

ud (difference) = u(model) - u (measured)
(where u is the average and the work done for each subject).

Ho : ud = 0   (u (model) = u (measured)).
H1 : ud <> 0   (u (model) <>u (measured)).

Hence the decision rule is reject Ho if
T(computed from the data) < The critical value 2.46
(alpha = .01).

The regression test was simply a way to gage the correlation between the actual and measured values. We plotted the model average vs the measured average for all subjects and did a linear regression on that data set. The same analysis was done on model vs measured work.

In addition to the above analysis, plots of actual vs measured raw torque values were also produced.

Results:

Figure 6 and 7 are regression plots of model vs measured averages (figure 6) and total work (figure 7). The correlation values (r = 0.854, and 0.842) indicate a strong relationship between measured and model values. This result indicates that the model can be used as a good predictor of the ratchet wrench torque produced when the model vs. measured values are compared for the entire subject pool in terms of the average torque produced and the total work done.

441

Pull operation - average torque for all subjects at two velocities

Push Operation - average torque for all subjects at two velocities

Pull operation - predicted vs measured work for two velocities

Push Operation - predicted vs measured work for two velocities
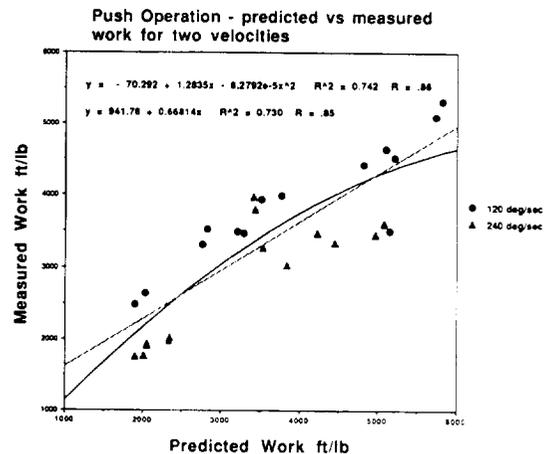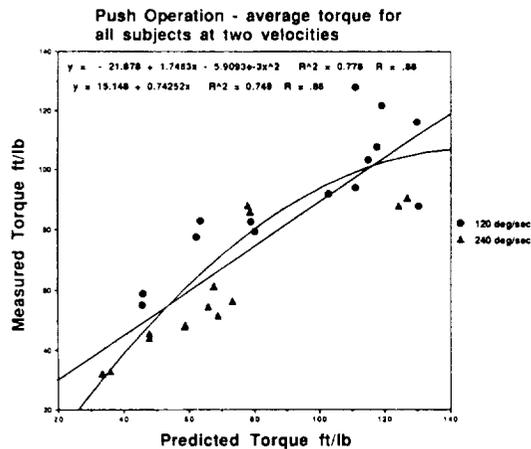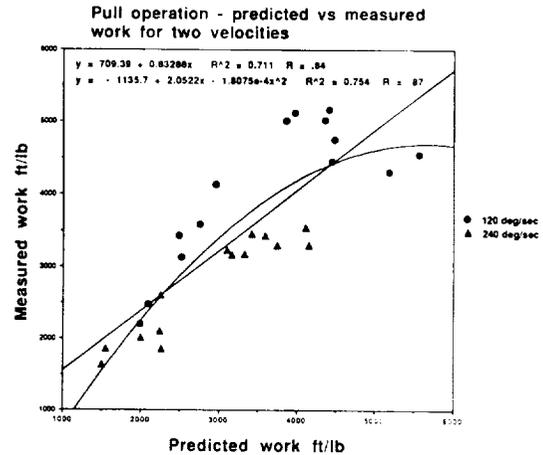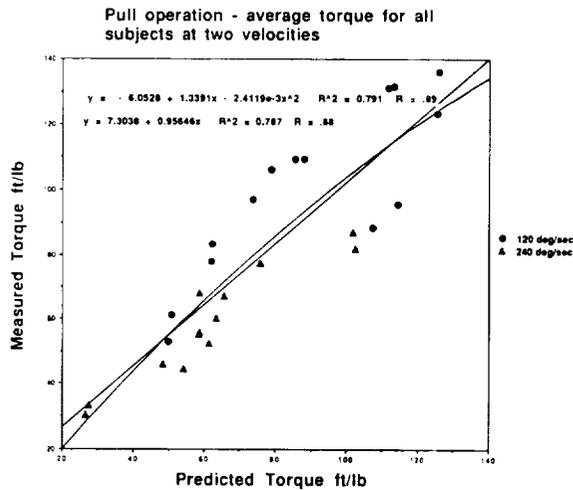
Figure 6 - Model vs measured average torque produced for the ratchet wrench motion.

Figure 7 - Model vs measured work produced for the ratchet wrench motion.

In addition to a regression comparison of the average and work done over the range of the motion, a pairwise T test is also performed on that data. Figure 8 is a table indicating T statistic results. These result indicate that for the ratcheting motion the model predicted and measured torque values show no statistical difference across the subjects at a level of alpha equal to 0.01.

Ratchet	Push:	Pull
Average torque	T = 1.96	T = -1.52
Total work	T = 1.96	T = -1.42

At alpha = 0.01 Critical value for degrees of freedom equal to 27 (14 subjects at 2 velocities -1) is 2.47.

ud (difference) = u(model) - u (measured) (where u is the average and the work).

Ho : ud = 0   (u (model) = u (measured)).
H1 : ud <> 0   (u (model) <>u (measured)).

Ho accepted because all T values calculated are within +/- 2.47, the critical value at alpha = 0.01.

Figure 8 - Pairwise T statistic results of average and total work across all subjects for all velocities.

Figure 9 shows a plot of the actual vs model torques for a male and female. This is an example set consisting of one male and one female. The two cases are extreme cases (strong male, weaker female). Trends in the data indicate that the model values are matched over five to 40 degree range of the ratchet motion. The initial and final few degrees (0-5, 40-45 degrees) do not match-up with our predicted results. These stages of the motion are related to the start-up and slow down processes involved [3] which are not currently part of our modeling effort. This result indicates that a similarity in magnitude and shape exists within subjects. Statistical tests are now in progress to study in detail the "goodness of fit" between the model and measured data for individual subjects in order to determine the limitations and enhancements to be made for the current model.

442
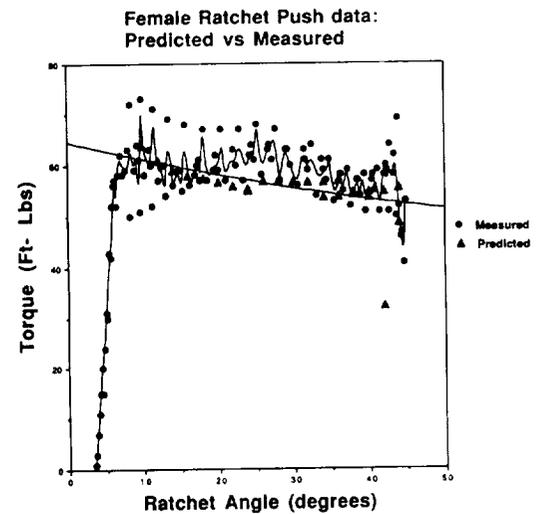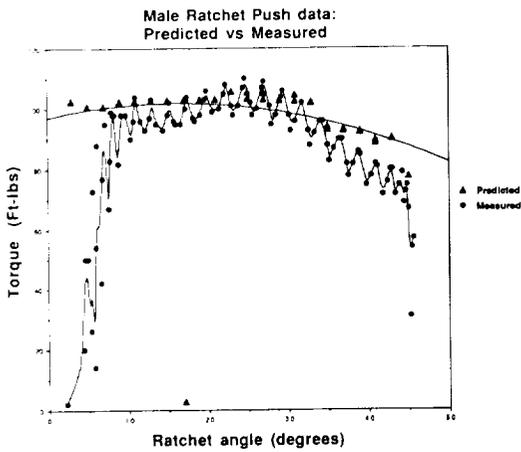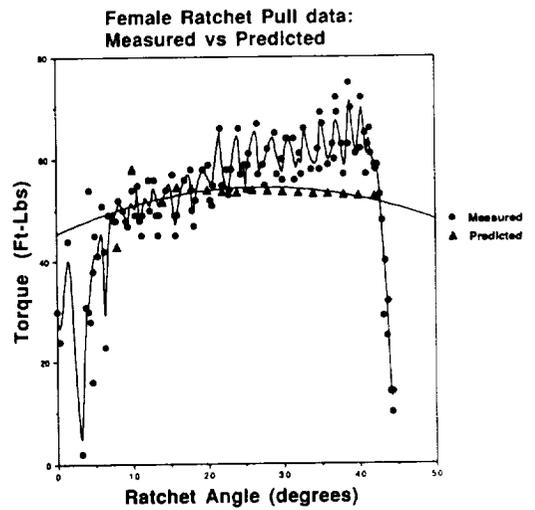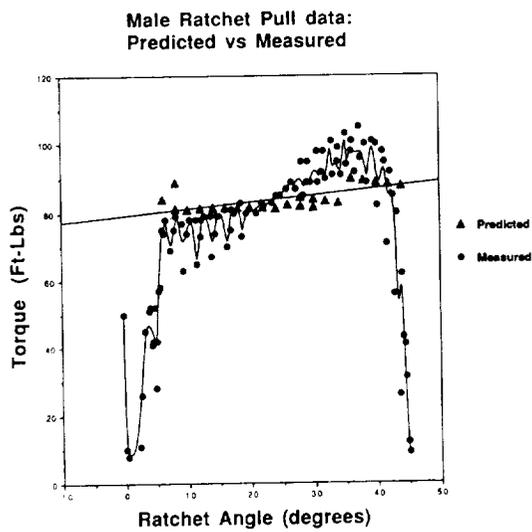
Figure 9: - Measured vs Model values for ratchet wrench maneuver.
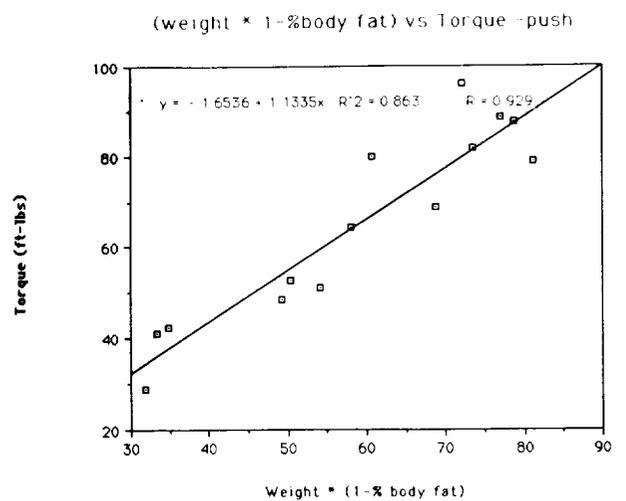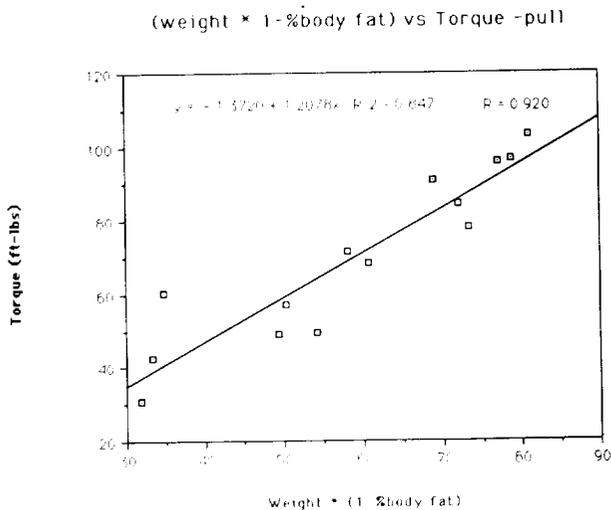
Figure 10- % Torque vs Weight (1- % body fat).

An important relationship exists which expands the utility of our model in terms of data collection requirements. A simple measurement, percentage body fat, is a good predictor of torque. Figure 10 is a plot of torque produced by all subjects in the ratchet wrench motion vs a calculation based on the body fat and weight (weight x (1-body fat)) for each subject. There is a strong correlation (r > 0.92) between torque production capability and the body fat calculation. Once a representative sample of a population has been measured for isolated joint strength, prediction of torque capability of a particular individual may be extrapolated by only two measure-- percentage body fat and weight. Research continues in this area.

Conclusions:

Unlike earlier attempts at strength modeling (based on rotational spring and damper systems) our model is based on empirical data. The shoulder, elbow, and wrist joints were characterized in terms of maximum isolated torque produced, position and velocity in all rotation planes for fourteen subjects. This information was reduced by least squares regression into polynomial equations relating torque produced as functions of position and velocity and tabularized for input to the strength model. This isolated joint information was used to compute (based on a vector sum algorithm and the subject's anthropometric measurements) forces resulting from composite motions-- in this case, the ratchet wrench push-pull. Measured vs model output were compared. (see figure 12)
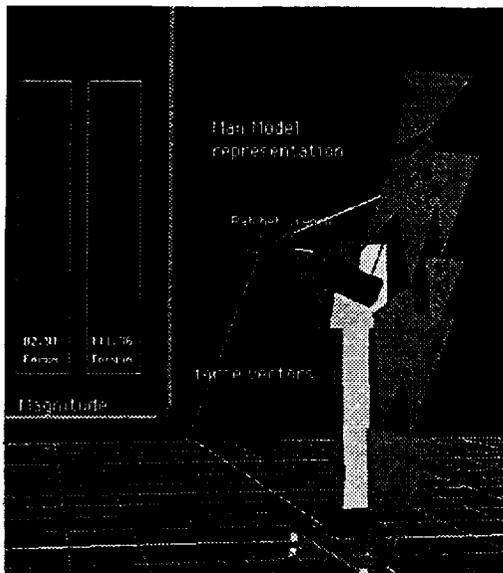
Figure 12 - Human model utilizing a ratcheting tool.

Results indicate that forces derived from a composite motion of joints (ratcheting) can be predicted from isolated joint measures. Model vs measured values for 14 subjects were compared. T values calculated were well within the statistically acceptable limits (alpha = 0.01) and regression analysis revealed coefficient of variation between actual and measured to be within 0.75 to 0.90 . Moreover, the model is flexible in terms of the environments and human motions that can be modeled. It has been demonstrated here that the current model predicts torque produced by a ratchet wrenching. Our overall objective is to incorporate into the existing CAE capabilities a strength model of the NASA crew member population for analytical human factors analysis. To this end, we will continue to cycle through the phases of validation and refinement with more complex motions and with additional isolated joint measures.

BIBLIOGRAPHY

1. Badler, N.I., Lee, P, Phillips. C. and Otani, E.M. "The JACK Interactive Human Model". In "Concurrent Engineering of Mechanical System Design in a Concurrent Engineering Environment, University of Iowa, Oct. 24-25, 1989.

2. Korein, James U. ' A Geometric Investigation of Reach', MIT Press, Cambridge, MA; 1985.

3. D.B. Chaffin and G.B. Anderson. *Occupational Biomechanics*. John Wiley & Son (New York), 1984.

4. C. Phillips, S. Wei, J. Zhau, N. Badler, " Strength Guided Motion", Computer Graphics 24(4):253-262.

5. F.T. Schanne, "Three Dimensional Hand Force Capability Model for a Seated Person" PhD Thesis, Univ. of Michigan, 1972.

6. NASA. *The Anthropometry Source Book*, Volume I and II. NASA Reference Publication 1024, Johnsons Space Center, Houston, Tx. 1978.

7. J Zhao and N.I. Badler. "Realtime Inverse Kinematics with Joint Limits and Spatial Constraints". Technical Report ms-cis-89-09, Dept of Computer and Information Science, Univ of Pennsylvania, Philadelphia, PA., 1989.

8. E. Rothman, W. Ericson, *Statistics: Methods and Application*, Kendall/Hunt Publishing (Iowa), 1987.

9. C. F. Gerald, *Applied Numerical Analysis*, Addison-Wesley Publishing Company (Menlo Park California) , 1970.

PERFORMANCE ASSESSMENT IN COMPLEX INDIVIDUAL AND TEAM TASKS

Douglas R. Eddy, Ph.D.
NTI, Incorporated
P.O. Box 35482
Brooks AFB Texas

ABSTRACT

This paper describes an eclectic, performance-based approach to assessing cognitive performance from multiple perspectives. The experience gained from assessing the effects of antihistamines and scenario difficulty on C^2 decision-making performance in Airborne Warning and Control Systems (AWACS) weapons director (WD) teams can serve as a model for realistic simulations in space operations. Emphasis is placed on the flexibility of measurement, hierarchical organization of measurement levels, data collection from multiple perspectives, and the difficulty of managing large amounts of data.

INTRODUCTION

Astronomers in the late 1700s recorded a star's transit by using a metronome to determine the moment a star touched the cross hairs of the telescope. The chief astronomer at Greenwich Observatory noticed that his assistant's times were consistently one second slower than his own. This was an early realization that the observer played a significant role in acquiring data and that even simple perceptual observations were susceptible to bias and individual abilities. Today we have sophisticated instruments to record much of the data of interest to science. However, in complex tasks where decisions must be based on human judgement or on the consensus of a team, the roles of the integrator of information and the decision-maker are still important and still susceptible to bias.

In the implementation of large projects such as building and maintaining a space station, building and maintaining a moon colony, or traveling to Mars, various designers need to know how our astronauts will handle the work. Engineers want to design consoles and workstations so operators can perform their tasks efficiently and without errors. Trainers want to provide timely and objective feedback to operators. Mission planners want to design work/rest cycles that maximize productivity while minimizing error and waste. Social planners want to provide work environments that facilitate team interaction and cooperation while minimizing the disruption of violations of personal space and privacy. The only way designers can have confidence about how our astronauts will perform on complex tasks far, far from home is by assessing performance in early design studies. Further, to maximize the use of equipment, facilities, subject time, and to obtain the most integrated data possible, complex, realistic, ground-based studies involving integrated payloads will be required. These future studies can benefit from the approach used in the Crew Performance Branch at Brooks Air Force Base, Texas, to assess antihistamine effects on complex task performance in WD teams and from lessons learned in the study. Since this paper is designed to communicate methods and approaches to understanding complex team tasks, emphasis is placed on experimental design issues with only sample results presented.

Although the primary goals of the study were to evaluate the effects of Seldane on complex performance, the researchers used the opportunity to gather data on several other issues from several perspectives. These included: the development of a methodology for assessing individual and team complex-task performance, the evaluation of sustained operations and fatigue, the assessment of cognitive workload through embedded tasks, the assessment of stress, the assessment of learning effects, the evaluation of tests for WD selection, and the prediction of complex task performance from cognitive skills tests.

THE WEAPONS DIRECTOR TASK

WDs in an air defense scenario must attend to a number of tasks. The wartime tasks include locating and identifying aircraft, maintaining track information on aircraft and targets, updating target information received from pilots, accepting aircraft hand-offs, performing a tactical controller function with appropriate level of control, providing target briefings to interceptors, performing a tanker controller function, providing recovery assistance, safe passage monitoring, briefing the senior director (SD) of any tracking or sensor data problems, and responding to alerts, alarms, and messages on the console. The success of the C^2 mission results directly from the WDs' successful accomplishment of their duties.

PERFORMANCE HIERARCHY

It is obvious that the performance of such a complex system including human operators is a result of numerous interacting internal and external factors. Because of these multiple determinants and numerous data perspectives, it was necessary to use a variety of measures to characterize the system and to diagnose the sources of observed variations in system performance. The interpretation of large metric sets is facilitated by an implicit underlying structure that weights the significance of each measure and relates it to the others.

After a review of the literature on objective measures of team performance (Eddy, 1989), the measurement aggregation problem was approached by devising a hierarchy of performance determinants that provides a classification framework for individual measures. Each level of the hierarchy contains groups of measures that jointly determine the measures available at the next level higher in the framework. For the command and control air defense scenario, four levels were chosen as shown in Figure 1. They are: Mission Effectiveness, System/team Performance, Individual Performance, and Performance Capability.
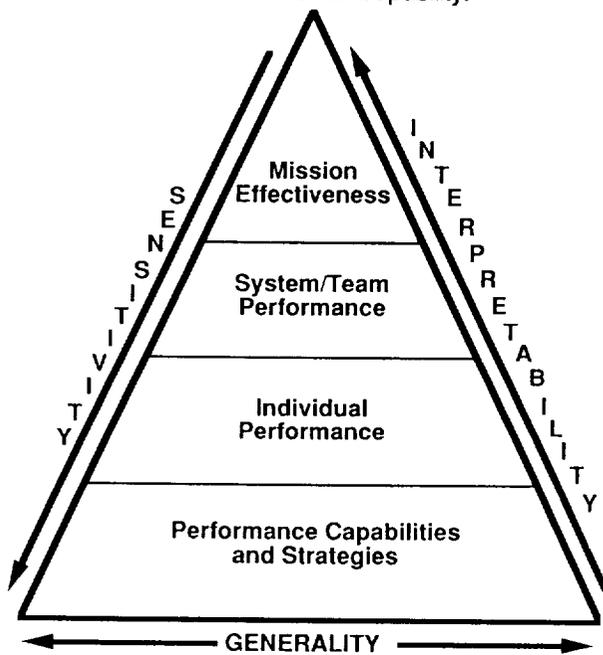
Figure 1. Performance measurement hierarchy.

The highest level of the hierarchy contains indices of Mission Effectiveness. These measures are derived directly from the specific objectives of the mission assigned to the system. An example is the protection of a specific sector of air and ground space from infiltration by enemy aircraft (protection of assets). Measures that flow from this objective and that assess performance in terms of mission effectiveness include the number of enemy infiltrations, the amount of fuel and weapons expended, and the ratio of enemy lost to friendly assets lost.

The second level of the hierarchy, System Performance, contains those groups of measures that reflect factors immediately affecting mission effectiveness. Such measures of System Performance reflect the degree to which the combined man-machine system has accomplished those tasks required to meet mission objectives. These measures do not reflect the individual contributions of different human behaviors or various hardware and software component performances. Instead, they are more global indices of the degree to which the total system successfully accomplished the tasks essential to mission success. For example, the weapons director/workstation system is required to meet its mission objectives essentially by accomplishing an air traffic control task aimed at directing interceptor aircraft to defeat threat aircraft. This air traffic control task decomposes into a number of essential subtasks such as pairing of interceptors with targets and providing target data to interceptors. A performance measure of the latter is the average accuracy and speed of data transfer to interceptor pilots.

The third level of the hierarchy is comprised of specific groups of measures that assess the individual contributions of human components to overall system performance. Measures included in the Human Performance level of the hierarchy are designed to reflect the quality of the individual behaviors required of the WD expressed primarily in terms of latencies and errors. These measures are derived by examining the system functions required to meet mission objectives in order to identify the specific contributions of the operator. For example, the system performance requirement to pair targets with interceptors requires the weapons director to identify a target's location on the workstation display and communicate this information to an interceptor aircraft via radio. The quality of the operator's performance in achieving this objective might be measured by evaluating the time needed to complete the full sequence of required behaviors and by assessing the accuracy of each manual and verbal response.

The final level of the proposed hierarchy contains measures that assess factors directly affecting the individual performance capacities of primary system components. For the human operator, measures of Performance Capability are composed of a large group of potential human state and ability measures that combine to determine overt performance. These measures include indices of workload or reserve processing capacity; fatigue; arousal level; experience level; and individual perceptual, cognitive, and motor abilities. This level also includes personality traits and predisposition to interact with teammates in specific ways that may or may not be adaptive under stress.

The multi-level classification of performance measures proposed above has the advantage of placing the measures into logical subordinate and superordinate groups indicating the predictive relationships among them. In addition, measures at each of the levels differ

446

in their sensitivity, generalizability, and practical interpretability.

REALISTIC SCENARIOS WITH EXPERIMENTAL CONTROL

Seldane, an antihistamine used in the study, does not cross the blood brain barrier and therefore may not affect performance. This puts the researcher in the unenviable position of trying to prove the null hypothesis. As a result it was necessary to demonstrate the sensitivity of the performance measures to some degrading treatment. Another antihistamine, Benadryl, was chosen for this purpose. (Merrell Dow Pharmaceuticals, Inc., provided the drugs for the study.) A placebo control brought the number of groups to three. Because it was possible to randomly assign experienced teams to one group and four inexperienced teams to another, we decided to collect data from each team under both difficulty conditions and under one antihistamine and placebo. Further, we wanted to collect two days of data under the drug conditions. Meeting these constraints required the development of six equivalent scenarios, but we had to prevent the subjects from anticipating the events in each succeeding scenario.

For the first scenario, we defined all the tracks, enemy flight paths, and events. We made lists of events and used a subject matter expert to indicate the impact on WD behavior of each. To prevent subjects from anticipating scenario events, in scenarios two through six we rotated the original so the enemy would appear from a different compass heading. Equivalent events were then spread across each scenario at the same points in time. We also changed land masses using different geographic locations, we created six unique prebriefings containing different political situations, countries, airbases, squadrons, call signs, and numbers. Debriefings at the end of all testing did not reveal that subjects believed any of the scenarios to be similar.

Scenario difficulty was manipulated in several ways. Enemy aircraft flew at varying altitudes and some took zigzag paths. The fog of war was increased by additional distractor events. Three scenarios were created for low difficulty and three for high difficulty.

METHODS

The 552d Air Wing assigned twelve teams of three WDs (male and female), who previously volunteered, to Brooks AFB to spend their work week in support of this study. The teams were randomly assigned to one of three drug treatment conditions and one of two scenario difficulty orders, either low-high or high-low.

The WDs arrived at Brooks AFB on either Saturday or Sunday evening for a preliminary briefing. Training took place on Monday for approximately eight hours. Teams received training on the AWACS-Performance Assessment Battery (AWACS-PAB), six simple computerized tests and two complex tests, over approximately four hours. The two complex tests were taken from the Complex Cognitive Assessment Battery that consists of nine tests.

The six tests were taken from Unified Triservice Cognitive Performance Assessment Battery (UTC-PAB) with over 25 tests. Further information on these tests can be obtained in Perez, Masline, Ramsey, and Urban (1987) and Hartel (1988). They also completed a three-hour C^3 training scenario to familiarize them with the simulated AWACS crewstations and scenarios. Subjects ingested one Benadryl and one Seldane placebo at 2230 or prior to going to sleep.

Starting on Tuesday, teams were then tested in two $3\frac{1}{2}$ hour scenarios each day for three days. Each group ingested only placebos during the testing schedule for Tuesday. A randomly assigned team ingested the recommended therapeutic dose of either Benadryl, Seldane, or a lactose placebo starting on Tuesday evening. Total antihistamine/placebo ingestion for each group across two days consisted of either eight 25mg Benadryl, four 60mg Seldane, or all placebo preparations.

One SD was used for all teams. After the prebriefing, his interaction with the team was to give direction only when required, but to keep the team from straying outside the performance measurement envelope. Other details of the facilities, equipment, scenario development and time schedules may be found in Schiflett, Strome, Eddy, and Dalrymple (1990).

Because the cognitive performance of the weapons director teams can be interpreted for a variety of questions, several subject trait, experience, and state measures were recorded. These included: a biographical sketch, a WD experience form, personality scales for potential use in developing WD selection tools, and surveys of their current state (symptoms, sleepiness, fatigue, etc.). The scales included the Rotter Scale, which assesses the locus of control generally perceived by a person in causing changes to take place in one's life; the Personal Characteristics Inventory (PCI), which assesses attitudes and leadership qualities; the Life Style Questionnaire, which predicts a subject's performance under stress; the Least Preferred Co-worker Scale (LPC), which may identify a WD's leadership style; the Jenkins Activity Scale, which assesses a WD's personality characteristics of decision-making; and the FIRO-B, which measures a subject's attitudes with regard to sociability and social interaction.

WD ratings were also obtained on the USAFSAM Fatigue Scale, which allows the subject to describe how he/she feels at that time; an Operational Impact Survey, which allows a subject to rate how well he/she felt the team completed its mission and how well each subject felt he/she completed his/her part of the mission; a Scenario Evaluation form, which allowed each WD to order the simulations with respect to difficulty; and the Subject Workload Assessment Technique (SWAT), which allowed each subject, at the

end of each simulation, to evaluate the workload of the scenario along SWAT's three dimensions: time load, mental effort, and psychological stress. The WDs kept logs similar to those kept during a standard mission. They recorded aircraft call signs, type aircraft, target numbers paired against, check-in time, weapons states on the aircraft at RTB, results, and other information.

In addition to the outcome measures of how well a team or individual is performing in a simulated air defense scenario, one would like to understand the underlying processes that contribute to those outcomes. Embedded tasks were used to measure *reserve capacity*, team coordination, and *situational awareness* (SA). These are tasks natural to the air defense scenario, but low priority. These tasks were delivered auditorially by voice queries articulated by the Votan speech synthesizer or by the SD.

The embedded measures for reserve capacity are: 1. whether or not a response is given, 2. accuracy of the response, and 3. latency of the response. The independent variables that may determine the WDs' workload level are: the number of flights currently under the WD's control, the level of control of each flight, the ADWL, and the number and type of additional tasks currently being worked by the WD. A typical SD query for reserve capacity might be "What state armament/fuel on the aircraft under your control?" Low difficulty should result in quick, accurate responses from the WD. High difficulty should result in ignored requests, partial information, and long response times.

Individual members of a WD team can work independently of each other. However, since the enemy is directing the attack in an air defense scenario, the battle does not always unfold the way it is planned in a mission prebriefing. As a result, each WD's responsibilities change throughout the mission. These changes should be adaptive and result from insight and leadership. Further, the adaptations require cooperation and coordination among the team members. WD responses involve passing and confirming information to each other and accepting responsibility for incoming requests when time is available. Embedded measures for team coordination include: 1. Whether or not the information is passed to the other WDs, 2. Accuracy of the response, and 3. Latency of the response. An event designed to elicit a team coordination response might be an ADWL announcement from ground control.

To effectively deal with events in an air defense scenario, a WD must maintain an accurate representation of the battle. This representation (both internal memory and external notes) defines the WD's awareness of the current situation. If the representation is in error, the WD may commit to kill rather than identify an unknown target. Therefore, throughout the scenario the WD's awareness was probed to determine if he/she has the correct ADWL, has kept track of airbase openings/closings, and tracked SAM sites going hot/cold. The embedded

measures for situational awareness are the same as for difficulty. An event designed to elicit a response would be for the SD to tell WD1 to kill track 0304. The WD should question this command during peace time since the SD had no authority to issue the order.

RESULTS

Early results for Mission and System/team performance levels and Performance Capability level have shown differential sensitivity to drug effects. At the Mission Effectiveness and System/team levels, 33 dependent measures, were amenable to statistical analysis. Of these measures, 6 showed a scenario difficulty effect, 4 showed a learning effect (days), and 8 showed a day by difficulty interaction. Table I shows the enemy penetrations, "get throughs," by day and difficulty. Although this variable did not achieve statistically significant results, it dramatically shows the impact of scenario difficulty and of performance improvement across days.

Table I. Enemy penetrations by day and scenario difficulty for all teams.

Condition	Penetrations
Day 2	
High Difficulty	22
Low Difficulty	5
Day 3	
High Difficulty	13
Low Difficulty	3
Day 4	
High Difficulty	6
Low Difficulty	6

Figure 2 shows the effect of scenario difficulty on the loss ratio of enemy/friendly aircraft. Loss ratios remained the same across days while ratios improved across days under low difficulty. In no case did performance under either antihistamine differ from the placebo group. These performance results for scenario difficulty were supported with WD ratings using the Subjective Workload Assessment Technique (SWAT). Generally scenarios designed for high difficulty resulted in higher workload ratings than those designed for low difficulty. A full description of the results can be found in Eddy, Dalrymple, and Schiflett (in preparation).

448
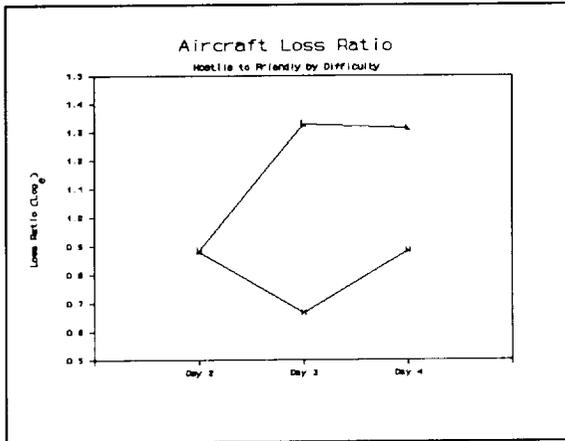
Figure 2. Effect of scenario difficulty on aircraft loss ratio.

By rearranging the data, time-of-day effects could be analyzed. Interestingly, performance on the morning simulations did not differ from that in the evenings, with difficulty balanced, even though subjective fatigue measures were higher during the evening simulation.

Benadryl degraded performance on cognitive skills and abilities as measured by the AWACS-PAB, especially on the first day of Benadryl administration, day 3 (Nesthus, Schiflett, Eddy, Whitmore, in preparation). Six of the tests, eight of the dependent measures, showed either a significant drug and/or drug-by-day effect. For example, Figure 3 shows an increase in errors in the Benadryl group on the Dichotic Listening test. Figure 4 shows that the Benadryl group found fewer word solutions on medicated days than the other groups who improved their performance on their treatment days. In addition, the Benadryl subjects subjective assessment of fatigue was greater on day 3. Seldane had no effect on performance as measured by these tests.
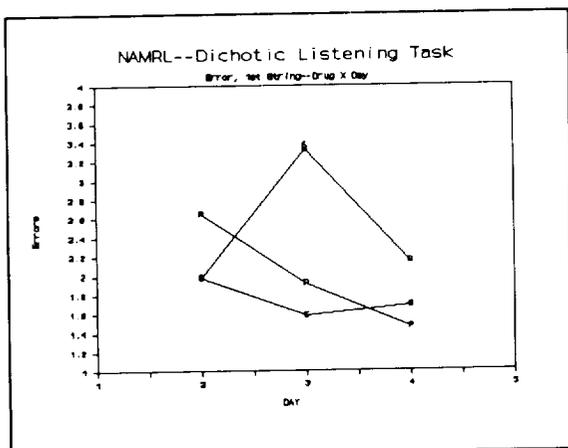
Figure 3. Antihistamine effects on dichotic listening.
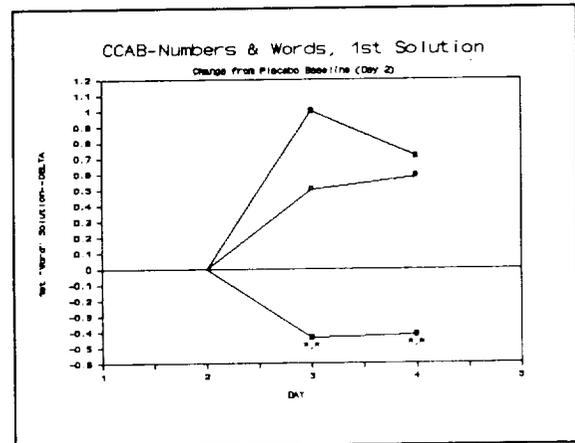
Figure 4. Antihistamine effects on the accuracy of the first solution in the Numbers and Words test.

FUTURE DIRECTIONS FOR THE AWACS DATA

The next step in data analysis involves developing rules for assigning individual WD responsibility within each scenario. These rules or definitions of areas of responsibility follow from the WDs training and practice. Once developed, each individual's role in "winning the war" can be assessed. This will include how well a WD controls his or her own area of responsibility (AOR), how he or she assists others, and how he or she requests assistance from the WD team. Through this approach the team's performance can be understood as a combination of individual efforts that either support or block the attainment of team goals. After the outcome measures of individual performance are obtained, process measures on the WD tasks and subtasks that produce the outcomes will be assessed. These measures will assess how well the individuals and teams accomplish such tasks as committing interceptors to targets, passing information to pilots, conducting intercepts, maintaining coverage of CAP points, maintaining situational awareness, etc.

MEASUREMENT PROBLEMS IN COMPLEX TASKS

To answer the question of why outcome measures of a system or team are degraded or improved, one must plunge into task analysis and modelling. Once good models of WD and team processes are established, objective measures of those processes can be evaluated against criterion individual and team outcome measures. Currently programmers and researchers in the AESOP laboratory are reviewing individual outcome measures and decomposing WD tasks and processes.

449

Several problems arise in attempting to objectively measure complex processes. For example, the beginning and ending of a task or process may not have well defined criteria or may cross media boundaries. We have often found that we can start with something concrete in the process, such as a switch action, and then work forward and backward for the start and end of the process. Sometimes this involves locating the switch action in the data file, obtaining the time stamp, using the time stamp to search through a file of transcribed utterances, and finally locating the initiating and/or ending event. This is a labor intensive process, but has the potential of being automated in the future with a text parser.

Another problem that arises involves simultaneous or overlapping tasks. Identifying when this happens and analyzing the single and dual tasks separately is one solution. If the same tasks overlap frequently and one task has a low priority, it may be possible to use the low priority task as an embedded secondary task to assess reserve capacity. Important tasks and processes that occur infrequently can provide highly variable latencies. If these tasks and processes have similar effects on the WD's behavior it may be possible to collapse the latencies of several treating them as a group. Often one task will interrupt another. This is an opportunity to verify the subject's prioritization of these tasks and if enough data exits, a confusability matrix can be generated.

LESSONS LEARNED FOR SPACE OPERATIONS

Because of the needs mentioned at the beginning of the paper, greater emphasis will be placed on understanding the effects of individual components on the performance of a complex system. This in turn calls out for the conduct of experiments with integrated payloads and performance measures to answer questions from multiple perspectives. As researchers, we must meet these needs by developing methods to assess performance in complex tasks. Our research on AWACS WDs has demonstrated that errors, failures, breakdowns in procedures, and systems may not show up unless the system is stressed. Researchers in space operations must continually search for system stressors that are realistic and appropriate to test a system's performance and its components. In this regard, statistical designs with repeated measures will be necessary to reduce variability, thereby requiring sophisticated ways of preventing subjects from anticipating events in repeated scenarios.

REFERENCES

Eddy, D. E. (1989). Selected team performance measures in a C³ Environment--An Annotated Bibliography. Technical Report USAFSAM-TR-87-25, USAF School of Aerospace Medicine, Brooks AFB, Texas.

Hartel, C. (1988). Expanded complex cognitive assessment battery (CCAB): Test descriptions. AAC-UM-33221, Systems Research Laboratories, U.S. Army Research Institute, VA: Alexandria.

Nesthus, T. E., Schiflett, S. G., Eddy, D. R., and Whitmore, J. N. (in press). Comparative effects of antihistamines on aircrew performance of simple and complex tasks under sustained operations.

Perez, W. A., Masline, P. J., Ramsey, E. G., and Urban, K. E. (1987). Unified tri-services cognitive performance assessment battery: Review and methodology. Technical Report AAMRL-TR-87-007, Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio.

Schiflett, S. G., Strome, D. S., Eddy, D. R., and Dalrymple, M. A. (1990). Aircrew evaluation sustained operations performance (AESOP): A Triservice Facility for Technology Transition. Technical Paper USAFSAM-TP-90-26, USAF School of Aerospace Medicine, Brooks AFB, Texas.

450

I

STRUCTURED ANALYSIS AND MODELING OF COMPLEX SYSTEMS

David R. Strome, Ph.D.
Systems Research Laboratories, Inc.
P.O. Box 35482
Brooks Air Force Base, Texas 78235

Mathieu A. Dalrymple
Systems Research Laboratories, Inc.
P.O. Box 35482
Brooks Air Force Base, Texas 78235

ABSTRACT

The Aircrew Evaluation Sustained Operations Performance (AESOP) facility at Brooks AFB, Texas, combines the realism of an operational environment with the control of a research laboratory. In recent studies we collected extensive data from Airborne Warning and Control Systems (AWACS) Weapons Directors subjected to low and high workload Defensive Counter Air scenarios. A critical and complex task in this environment involves committing a friendly fighter against a hostile fighter. Structured Analysis and Design techniques and computer modeling systems have been applied to this task as tools for analyzing subject performance and workload. This technology is being transferred to the Man-Systems Division of NASA/JSC for application to complex mission-related tasks, such as manipulating the shuttle grappler arm.

INTRODUCTION

Structured analysis and modeling are not new tools. They have been used for many years as aids in defining and analyzing systems, projects, products, and concepts. This paper discusses one application of these tools to a highly complex Air Force operational task, and the transfer of this technology to the Man-Systems Division of NASA.

Modeling focuses our attention on the processes and relationships of a system. It allows us both to describe a system as it is and to predict system behavior when conditions or constraints are altered. The accuracy of the prediction depends on the complexity of the question and the detail and validity of the model. Although often thought of as a research method, computer modeling is used in many forms: financial systems are modeled with spreadsheets; projects are modeled with program management software; motion is modeled with specialized graphics systems. Research models use specialized programs to explore neural function,communications, strategies and tactics, human performance, and task loading. These are only a few of the multitude of applications.

Creating a model is a simple concept but a demanding process. Using input from experts, the system being modeled must be broken down into detailed steps or processes, the relationships between the steps must be clearly defined, and the parameters that affect each process or relationship must be determined. The model is strengthened, verified, and validated by testing the real world system under a variety of conditions and comparing the data with the model's results. This iterative procedure continues until the model reaches a level of description and prediction that meets the demands of the research.

Structured analysis is a formalized process for developing the detailed information required to build the model. As a process, it is accomplished when any model is built, whether it is done formally or not. For simple systems models can be adequately defined with less rigorous methods. Complex tasks, on the other hand, benefit from a more structured technique that aids the analysis. Structured Analysis and Design Technique (SADT) [Marca and McGowan, 1988] is a widely recognized implementation of this concept. Originally a paper and pencil exercise, it is now appearing in software tools that provide basic consistency checks, easy modification, and presentation output. The basic building block of SADT is a box containing a descriptor of an event or process. This event has inputs, outputs, controls, and resources or mechanisms (Figure 1).

(controls)

(inputs) | Event | (outputs)

(resources/mechanisms)
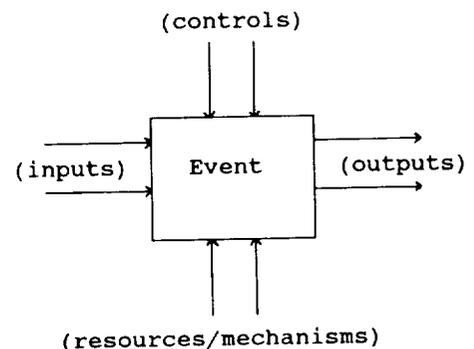
Figure 1

These blocks are linked together hierarchically in increasing detail until the model is sufficiently described. Structured analysis:

+ improves system definition.

+ improves problem understanding.

+ improves user/developer communication.

+ impacts system design.

+ smooths transition from analysis to design.

Technical developments in computer hardware and software have supported attempts to marry these two tools into a single integrated package of analysis and modeling. As separate tools, the information entered into the structured analysis program has to be re-entered into the modeling software. To overcome this duplication of effort, software engineers are developing integrated systems that automatically port the structured information into selected modeling systems.

AIR FORCE APPLICATION

The Aircrew Evaluation Sustained Operations Performance (AESOP) facility is located in the Sustained Operations Branch of the Crew Technology Division at Brooks Air Force Base, San Antonio, Texas. It is a research facility dedicated to the following goals:

+ Develop and apply individual and team performance measures to evaluate crew interactions under sustained operations.

+ Develop and integrate tools to support research and development.

+ Transfer technology to operational environments.

During 1989 the Air Force conducted a study in AESOP using 12 teams of Airborne Warning and Control Systems (AWACS) Weapons Directors (WDs). WDs provide control for Air Forces in their area of responsibility, committing friendly aircraft to missions such as: identification of unknown aircraft; engagement of hostile aircraft; search and rescue operations; mid-air refueling; and escort. Their operational environment is a complex combination of tactics and strategy, decision-making, and communication.

Each team of three WDs was subjected to three high workload and three low workload defensive counter air scenarios, each lasting three and one-half hours. Among the collected operational data were: switch actions, key presses, communication channel usage, audio traffic, microphone activations, target locations, and mission critical events. Video tapes captured non-verbal, non-console subject interaction. In addition to

operational measures, a battery of individual performance tests, surveys, and questionnaires was given throughout the course of the study. The resulting data is being continually analyzed to develop and test individual and team performance measures directly related to operational tasks.

Part of the ongoing development of tools to assist in this research involved selecting a complex WD task for structured analysis and modeling. The purpose was to evaluate several software systems while determining the applicability of the tools to our environment. The data from real subjects provided a testbed for the validity of the model.

Committing a friendly fighter to a mission against a potentially hostile, or known hostile, target is a primary WD task. The first two levels of the SADT analysis, created using Idefine (Wizdom Systems, Inc.), are shown below (Figures 2,3)
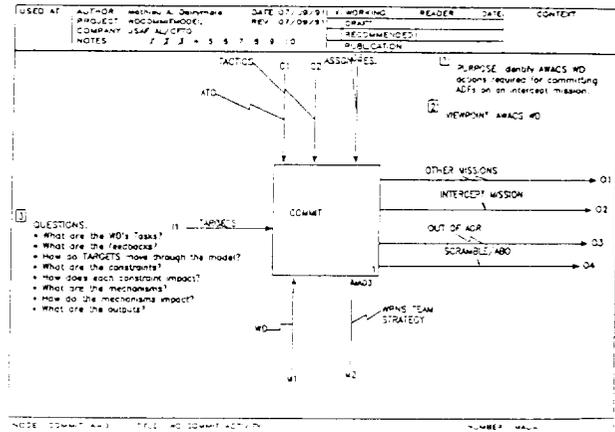
Figure 2

At the top level (Figure 2) we define the broadest description of the task. *Targets* on the computerized radar screen are the INPUTS to the task. The *WDs* and their *team strategy* are MECHANISMS for completing the task. The *Air Tasking Order (ATO)*, *tactics*, and *assigned resources* are the CONTROLS (CONSTRAINTS) on conducting the task. When the task is complete the possible OUTPUTS include a target being *out of Area Of Responsibility*, a *scramble/airborne order* requesting additional resources, a commit of existing resources to an *intercept mission*, or committing to some *other mission* such as search and rescue or escort. The next part of the hierarchy (Figure 3) divides this broad analysis into the next level of detail.

This level has the same overall inputs, controls, mechanisms, and outputs; but the task is now subdivided into *identifying targets, sorting threats, sorting friendlies,* and *committing to a mission*. These tasks are connected functionally with outputs becoming inputs or controls to other steps as necessary. Each of these boxes is then subdivided in the same manner at the next
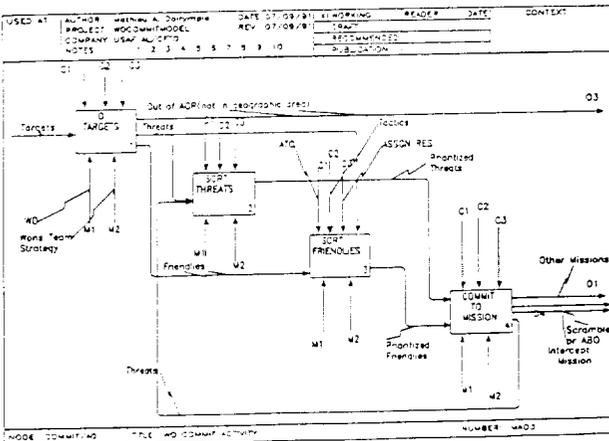
Figure 3

level of the hierarchy, and so on until reaching the required depth of detail.

The finest level of detail is entered into the modeling software. Time to completion, probability of a particular path, and other selected parameters are defined for each task to complete the definition of the model. Figure 4 is a typical output, graphing the frequency distribution of time required to complete the commit task.
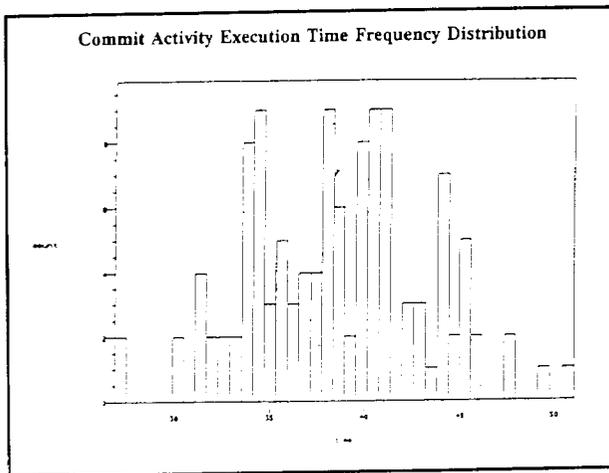
Figure 4

The average completion time is around 37 seconds, which matches well the average time collected in actual scenarios. Changes in the model result in predictable shifts in this average time. However several aspects of the performance of real WDs are not accounted for by the current model. The model tends to have much less variation than human operators. It also does not account well for very slow or very fast times, which may be due to inattention, fatigue, anticipation, distraction, or other external causes. Finally it does not partition realistically between component phases of the overall task, such as a decision phase where the WD is

deciding what to do, and the action phase where switches and keys are pressed and information is communicated. Our experience suggests that the accuracy required to adequately describe and predict real world behavior will demand that these aspects be accounted for. To do so will require models that provide true parallel processing and accept algorithms for dynamics such as fatigue, anticipation, and distraction.

TECHNOLOGY TRANSFER

Working under the guidelines of a memorandum of understanding, the Sustained Operations Branch is working with the Crew Interface Analysis Section of the Man Systems Division at NASA/Johnson Space Center to integrate these tools into the space operations and research environment. The first task to be analyzed and modeled by NASA is the operation of the Remote Manipulator System at both normal and reduced gravity.

CONCLUSION

The application of these tools to both Air Force and NASA operational tasks has emphasized their usefulness in defining and understanding complex systems. Future use of structured analysis and modeling in the AESOP will aid in:

- ◆ Determining critical performance elements in complex tasks.

- ◆ Predicting performance effects of fatigue, decision aids, and drugs.

- ◆ Defining and testing new training strategies.

- ◆ Developing meaningful individual and team performance measures.

REFERENCES

Marca, David and McGowan, Clement, *SADT Structured Analysis and Design Technique*, McGraw Hill, New York, NY, 1988.

453

Evaluation of Force-torque Displays for Use with Space Station Telerobotic Activities

Robert C. Hendrich[1], John M. Bierschwale[1], Meera K. Manahan[1], Mark A. Stuart[1], and A. Jay Legendre[2]

[1] Lockheed Engineering
and Sciences Company
2400 NASA Rd. 1 C95
Houston, TX 77058

[2] NASA - JSC
SP34
Houston, TX 77058

ABSTRACT

Recent experiments (NASA/MSD, 1991) which addressed Space Station remote manipulation tasks have found that tactile force feedback (reflecting forces and torques encountered at the end-effector through the manipulator hand controller) does not improve performance significantly. Subjective responses from astronaut and non-astronaut test subjects indicated that force information, provided visually, could be useful. No research exists which specifically investigates methods of presenting force-torque information visually. This experiment was designed to evaluate seven different visual force-torque displays which were found in an informal telephone survey. The displays were prototyped in the HyperCard programming environment. In a within-subjects experiment, fourteen subjects nullified forces and torques presented statically, using response buttons located at the bottom of the screen. Dependent measures included questionnaire data, errors, and. response time. Subjective data generally demonstrate that subjects rated variations of pseudo-perspective displays consistently better than bar graph and digital displays. Subjects commented that the bar graph and digital displays could be used, but were not compatible with using hand controllers. Quantitative data show similar trends to the subjective data, except that the bar graph and digital displays both provided good performance, perhaps due to the mapping of response buttons to display elements. Results indicate that for this set of displays, the pseudo-perspective displays generally represent a more intuitive format for presenting force-torque information.

INTRODUCTION

Space Station *Freedom* will employ multiple telerobotic systems in its assembly and maintenance. These systems will have the capability to provide information back to the operator concerning the forces and torques encountered at the end effector (force feedback). Recent experiments at NASA's Johnson Space Center (JSC) (NASA/MSD, 1991) have found that for Space Station tasks, tactile force feedback provided through the manipulator hand controller (force reflection) does not improve performance significantly. However, qualitative responses from astronaut and experienced non-astronaut subjects indicate that force information can be useful. Force feedback can be provided via the visual modality (Hannaford, Wood, Guggisberg, McAffee, and Zak, 1989; Molino, Farbry, Langley, and Fisher, 1990) as well as the tactile modality (Hannaford, et al., 1989; Garcia, Chapel, and Spofford, 1990).

Bar graphs have been investigated as a means of presenting force-torque information visually (Bejczy & Paine, 1978; Bejczy & Dotson, 1982; Bejczy, Dotson, Brown, & Lewis, 1982; Molino, et al., 1990). Bejczy, et al., (1982b) found that the visual display aided the operators in performing a payload berthing task, especially in the terminal phase of berthing. Subjects in Molino, et al., (1990) commented that bar graphs were useful in situations in which the manipulator was bound up due to excessive contact forces. A pseudo-perspective graphic display of force-torque information (Figure 1, Display 1) was developed at the Jet Propulsion Laboratory (JPL). In this study, Corker, Bejczy and Rappaport (1986) found that the use of this display reduced the force

454

|
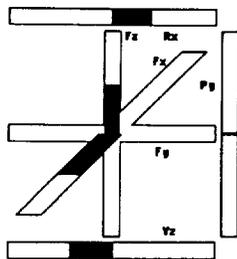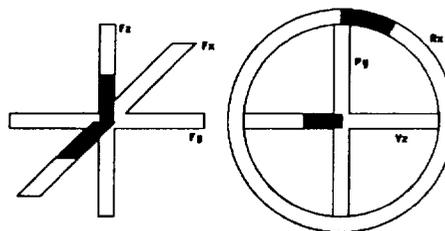
Display 1

Display 2
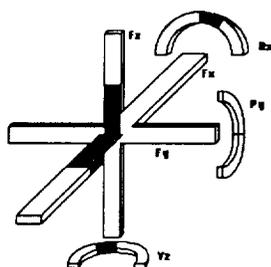
Display 3
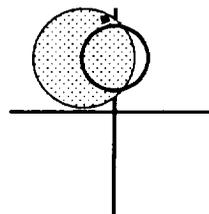
Display 4
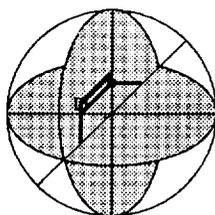
Display 5
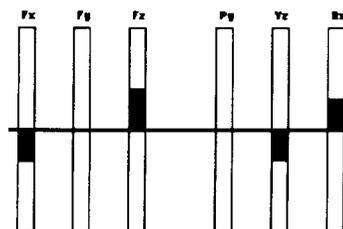
Display 6

Display 7

Figure 1. The seven displays investigated, each configured with equivalent forces in X, Z, Yaw, and Roll.

applied to payloads in a Remote Manipulator System (RMS) berthing task by 30-50% over the same task performed without the display. In a related study, Hannaford, et al., (1989) had one subject perform three different tasks using the JPL display and found that the visual display provided better performance than no feedback, but worse than tactile force feedback. In all of the visual force-feedback research previously conducted, none exists which directly investigates the best method of presenting force-torque information visually. Therefore, a need existed for an evaluation of this type.

An informal telephone survey was conducted to determine if there were displays that have been developed outside of JPL. Contacts were made with NASA personnel, universities, the nuclear industry, and aerospace contractors. Two alternative displays were found, one developed at Langley Research Center (Figure 1, Display 4) and a second (Figure 1, Display 5) developed within the Automation and Robotics Division at JSC. Two other displays (Figure 1, Displays 2 and 3) were developed by the Remote Operator Interaction Laboratory (ROIL) at JSC and are essentially variations of the JPL display. Display 6 is a generic bar graph display and Display 7 is a generic digital display included as a baseline.

This experiment was designed to evaluate the displays found in the survey in a controlled environment. The results may also be used to make modifications to the prototype displays and/or suggest guidelines for the display of force-torque information.

METHOD

Subjects/Experimental Design

Two groups of seven subjects participated. The first group consisted of volunteers from the Man Systems Division at JSC, all of whom had experience using a remote manipulator with 2 x 3 degree-of-freedom (DOF) hand controllers (separate controllers for rotation and translation, which are baselined for Space Station Freedom). The second group consisted of volunteers from JSC who had experience with the shuttle RMS simulator. While both groups had experience with remote manipulation, the Man Systems group was more experienced in human factors aspects of display design and the second group had more experience with operational concerns of using a remote manipulator.

Group (2), display (7), and number of axes with forces displayed (6) served as independent variables. Display and number of axes having force were within-subjects variables and group was a between subjects variable. The number of axes displaying force was controlled such that each display had four trials in which all six axes displayed forces, four trials with five axes of force, etc. Dependent measures included completion time, errors, and subjective data collected through questionnaires given after each display and at the completion of the experiment.

Apparatus

The experiment was conducted in the ROIL at JSC, using a Macintosh IIfx computer in the HyperCard programming environment. Both completion time and error data were collected by the program. The seven displays investigated are illustrated in Figure 1.

Procedure

The subject's task was to nullify the forces and torques presented by each display. This was accomplished via a set of buttons at the bottom of the screen which enabled the subject to manipulate the display in both positive and negative directions for each of the six axes (X, Y, Z, Pitch, Yaw, and Roll). If, for instance, a force was displayed in the positive X direction, the subject had to use the mouse and click the negative X button to eliminate the force. Each subject received 30 trials with each of the seven displays, with the order of display presentation being counterbalanced across subjects. The first six trials were practice trials which allowed subjects to become familiar with the operation of a display. After completion of 30 trials with a display, the subject filled out a questionnaire which addressed certain aspects of how well the display allowed the subject to detect the presence and monitor the changes of the displayed information, in addition to how well the display could be used with a set of 2 x 3 DOF hand controllers. After the subjects completed all seven displays, they were given a final questionnaire which allowed them to rate

456

the displays after having seen each one. For both types of questionnaires, subjects were encouraged to rate the displays based on strictly the *concept* behind the displays (how intuitive they were for presenting force-torque information), and not how well the displays worked with the response buttons nor how smoothly the displays moved on the screen.

RESULTS

Analyses of variance (ANOVA) were performed on both the quantitative data and the mean ratings from each question of the questionnaires. In addition, Tukey's test of paired comparisons was run with each ANOVA to determine differences among displays and groups. Data from the two groups of subjects showed no significant differences between the groups, either for the final questionnaire or post-display questionnaire data. In addition, no significant differences between the groups were found for either errors or completion time. Therefore, the data for the two groups were pooled and analyzed as one group of fourteen subjects. Also, data from the post-display questionnaires were very similar to the data from the final questionnaire and will not be presented here.

Subjective Data

Final questionnaire data were collected using seven-point scales (1 corresponding to "completely acceptable", 7 corresponding to "completely unacceptable"). For question 1, which involved how acceptable the displays were for presenting forces in X, Y and Z, there was a significant main effect of display, $F(6,78)=11.78$, $p<.001$. The test of paired comparisons showed that Display 5 was rated significantly worse than all other displays, with Display 7 rating significantly worse than Display 3. Displays 1, 2, 3, 4, and 6 did not show any significant differences among each other. The main effect of display was also significant for question 2, $F(6,78)=13.36$, $p<.001$, which concerned the acceptability of the displays for presenting torques in Pitch, Yaw, and Roll. The test of paired comparisons showed the same differences as for question 1, except that Display 7 was rated significantly different than Display 2 instead of Display 3. Question 3 asked how compatible the displays were with using 2 x 3 DOF hand controllers.

Again, a significant main effect of display was found, $F(6,78)=28.32$, $p<.001$. Displays 5, 6, and 7 all were found to rate significantly worse than Displays 1 through 4 in the test of paired comparison. Displays 5, 6, and 7 were not significantly different than each other. Finally, question 4, which involved how acceptable the displays were overall, demonstrated a significant main effect of display, $F(6,78)=13.71$, $p<.001$. The only difference in the test of paired comparisons showed Display 5 to be rated significantly worse than all other displays. Data from the final questionnaire are presented in Figure 2.

Quantitative Data

The main effect of display was significant for completion time $(F(6,108) = 13.22, p<.001)$.
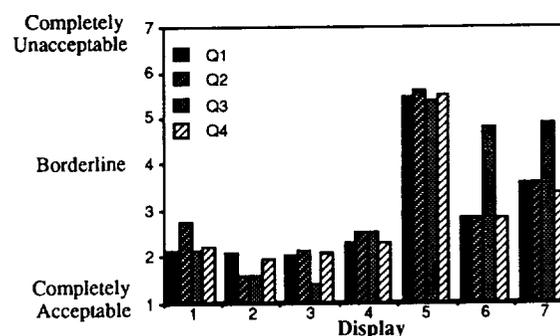
Figure 2. Mean ratings for each display by question.

The test of paired comparisons showed that Display 5 was significantly slower than every other display, while Display 4 was significantly slower than Displays 6, 7, and 2. Completion time data are shown in Figure 3.
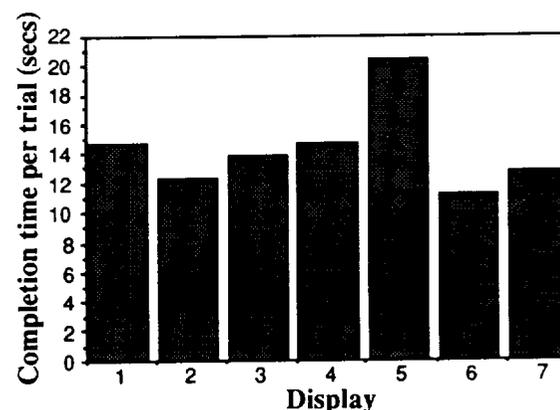
Figure 3. Mean completion times.

457

Three types of errors were recorded. Overshoot errors occurred when a subject, in eliminating a force or torque, went past zero and caused a force in the opposite direction. Reversal errors were scored when a subject applied more force instead of eliminating force. Confusion errors resulted when a subject attempted to zero-out a force that was already zero. For all three types of errors, the main effect of display was significant (overshoot errors, $F(6,108) = 7.69$, $p<.001$; reversal errors, $F(6,108) = 6.08$, $p<.001$; confusion errors, $F(6,108) = 5.34$, $p<.001$). Analysis of paired comparisons showed that Display 5 had significantly more overshoot errors than all other displays, and the remaining displays were not different than one another. For reversal errors, Display 5 had significantly more errors than all other displays, and Display 4 had significantly more errors than Display 7. Display 5 had significantly more confusion errors for all other displays as well, with Display 1 having significantly more errors than Display 6. Error data are shown in Figure 4.
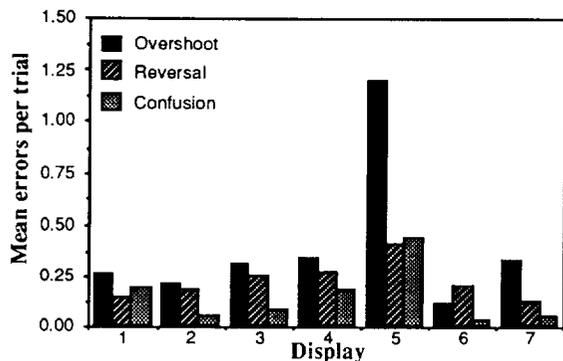
Figure 4. Mean overshoot, reversal, and confusion errors for each display.

DISCUSSION

Results from the final questionnaire can be broken down into two sections. Three of the four questions (1, 2 and 4) in the final questionnaire showed very similar relationships between the displays. For these questions, Displays 1 through 4 rated slightly better than Displays 6 and 7, with Display 5 being rated worst. However, for the question which asked about the compatibility of the displays with 2 x 3 DOF hand controllers (question 3), Displays 6 and 7 had similar ratings to Display 5, all of which were rated much worse than Displays 1

through 4. These poor ratings for Displays 6 and 7 were not unexpected, as there is no relationship between the elements of the displays and the spatial nature of force and torque information. Comments by subjects indicated that the method used by Display 5 of presenting torques, especially Roll, was not intuitive. In addition, when more than 3 forces or torques were displayed simultaneously, the intersecting lines created confusion as to which axes had forces. This situation caused several subjects to adopt a trial and error approach in identifying the displayed forces. A trial and error approach to relieving forces in an actual manipulation task, where highly expensive and delicate equipment is involved, is not a preferred strategy to adopt.

The results from the quantitative data differ from the subjective data mainly in the performance provided by Displays 6 and 7. Both displays had very few errors and generally faster completion times than all other displays. Their superior performance might be a result of an advantageous mapping of the display elements onto the response buttons. For both displays, the axes were presented from left to right on the screen as X, Y, Z, Pitch, Yaw, Roll, in the same order as the response buttons at the bottom of the screen. Due to the simplicity of the task, the number of errors committed was very low, generally less than one error per trial. It is interesting to note that overshoot errors followed the same pattern of results as the rest of the data, even though the occurrence of these errors was largely tied to how smoothly the displayed elements moved on the screen.

Display 5 consistently provided the poorest performance across all dependent measures. Display 4, was generally next-poorest to Display 5 with respect to the quantitative data. Display 4, however, was rated highly in both the post-display and final questionnaires. Subjects commented that this display was very intuitive in presenting torque information and was compatible with using 2x3 DOF hand controllers.

Displays 1, 2, and 3 had similar formats and produced similar ratings. Displays 2 and 3 however, generally rated slightly better due to their more intuitive methods of representing torques, as some subjects reported confusion

458

between the Yaw and Roll axes on Display 1. Several subjects commented that the torque information on Display 3 mapped very well onto movement of a hand controller. For all three displays a large number of subjects commented that the location where the three force axes intersect needs some kind of graphic which would help in differentiating small forces.

CONCLUSIONS

Among these seven displays, Displays 1 through 4 seem to represent the most intuitive formats for presenting force-torque information. The results indicate that the "standard" or most widely known display, the JPL display (Display 1), may be improved upon significantly by simply modifying how torque information is presented. Display 4 represents a completely different method of presenting the information, and further testing in a more realistic task environment is needed to clarify the differences and make recommendations for display selection. The task utilized here did not involve a working manipulator or hand controllers, and the displays were essentially presented statically.

A second experiment is planned in which the top candidate displays identified in this study are incorporated into a working robotic system, with a 6 DOF manipulator and a set of 2 x 3 DOF hand controllers, to perform a task in which forces and torques must be observed and controlled. The selection of a display format may depend on other considerations such as screen space and computing power available. Both the use of color and auditory cues were not addressed by this study, as they may be added to any of these displays in a similar fashion.

REFERENCES

Bejczy, A. & Paine, G. (1978). Event-driven displays for manipulator control. *Proceedings of the the 14th Annual Conference on Manual Control.* Los Angeles, CA: National Aeronautics and Space Administration.

Bejczy, A. & Dotson, R. (1982). A force-torque sensing and display system for large robot arms. *Proceedings of the IEEE Southeastcon.* Destin, FL: IEEE.

Bejczy, A., Dotson, R., Brown, J., & Lewis, J. (1982). Force-torque control experiments with the simulated shuttle manipulator in manual control mode. *Proceedings of the 18th Annual Conference on Manual Control.* Dayton, OH.

Corker, K., Bejczy, A. & Rapparport, B. (1985). Force/torque display for space teleoperation control experiments and evaluation. *Proceedings of the 21st Annual Conference on Manual Control.* Columbus, OH.

Garcia, K., Chapel, J., & Spofford, J. (1990). Effect of hand controller characteristics on operator performance. *International Symposium on Measurement and Control in Robotics.* Houston, TX: Clear Lake Council of Technical Societies.

Hannaford, B., Wood, L., Guggisberg, B., McAffee, K., & Zak, H. (1989). Performance evaluation of a six-axis generalized force-reflecting teleoperator. (JPL Publication 89-18). Pasadena, CA: Jet Propulsion Laboratory.

Molino, A., Farbry, J., Langley, L., & Fisher, D. (1990). Robot/operator performance on the work package 4 battery ORU changeout task. (Tech. Report TUF 90-04). Alexandria, VA: Tech-U-Fit Corporation.

Man Systems Division (1991). *Space station hand controller commonality test report.* (NASA Report JSC-32125). Houston, TX: National Aeronautics and Space Administration.

459

MEASUREMENT OF PERFORMANCE USING ACCELERATION CONTROL AND

PULSE CONTROL IN SIMULATED SPACECRAFT DOCKING OPERATIONS

Adam R. Brody
Sterling Software
NASA Ames Research Center
MS 262-2
Moffett Field, CA. 94035-1000

Stephen R. Ellis
NASA Ames Research Center
MS 262-2
Moffett Field, CA. 94035-1000

ABSTRACT

Nine commercial airline pilots served as test subjects in a study to compare acceleration control with pulse control in simulated spacecraft docking maneuvers. Simulated remote dockings of an orbital maneuvering vehicle (OMV) to a space station were initiated from 50, 100, and 150 meters along the station's -V-bar (minus velocity vector). All unsuccessful missions were reflown. Five-way mixed analyses of variance (ANOVA) with one between factor, first mode, and four within factors, mode, block, range, and trial were performed on the data. Recorded performance measures included mission duration, and fuel consumption along each of the three coordinate axes. Mission duration was lower with pulse mode while delta V (fuel consumption) was lower with acceleration mode. Subjects used more fuel to travel faster with pulse mode than with acceleration mode. Mission duration, delta V, X delta V, Y delta V, and Z delta V all increased with range. Subjects commanded the OMV to "fly" at faster rates from further distances. These higher average velocities were paid for with increased fuel consumption. Asymmetrical transfer was found in that the mode transitions could not be predicted solely from the mission duration main effect. More testing is advised to understand the manual control aspects of spaceflight maneuvers better.

INTRODUCTION

Historically, in the design of large and complex systems such as aircraft, automobiles, and nuclear power plants, designers typically ignored human factors considerations or left them until too late in the design process to be useful. Controls and displays located outside reach and sight envelopes, inappropriate automation, and operating procedures designed without concern for man-in-the-loop considerations have plagued various industries and led to the loss of many lives, vehicles, and other equipment. These accidents are highly visible in the aviation industry where two-thirds of the commercial aviation incidents and almost 90% of the general caused or influenced by human error.[1]

The Federal Aviation Administration (FAA), National Aeronautics and Space Administration (NASA), Department of Defense (DOD), and the major airline manufacturers are actively involved with investigating the human factors environment of aircraft to identify the means to reduce the likelihood of human error. "At first sight, it is a strange professional link between the aerospace design engineer and the psychologist. Yet, since the days of the Wright brothers, there has always been a need for designers to take human factors into consideration to ensure the efficiency of any flying machine."[2] Two particular concerns are automation and crew coordination and their relationships with flight procedures.

While spaceflight does not put millions of civilians at risk every day, every minor incident receives tremendous attention by the media and the public. On-orbit flight activities put lives, missions, and billions of dollars of hardware in jeopardy. Current and future research into the manual control aspects of orbital flight will have tremendous payoffs in safety, reliability, efficiency, and productivity as space traffic increases in the upcoming Space Station Freedom era.

Spacecraft docking will be a commonplace activity in the era of the space station. Shuttle orbiters, orbital maneuvering vehicles (OMV) (or equivalent), and orbital transfer vehicles (OTV) will be docking to the station. Vehicles will dock with satellites as well to return them to the station. Further into the future, vehicles will be docking in orbit around Mars, in lunar orbit, and on return to Earth orbit. Space Station Freedom will be used as a staging area for assembling and verifying spacecraft en route to the moon and Mars. It will also be used as a repair shop for satellites and a platform for experiments and equipment. These activities will increase the docking traffic at the station further justifying current research agendas. Current state-of-the-art computer graphics has improved real-time simulation and intensive and comprehensive human factors investigations with which researchers can study and better understand these activities.

Very little research describing human factors implications of spacecraft docking operations has been documented in the twenty-five years since the first spacecraft docking.[3-20] Parameters of flight such as approach and impact velocities, braking gates, and control modes must be examined to uncover fundamental human factors capabilities and shortcomings with regard to piloting spacecraft. Results from these studies will assist in expanding the operational flight envelope, and increasing safety and productivity. This study represents another in a series of experiments designed to accumulate a comprehensive database describing the manual control aspects of orbital flight.

Practical exploration by man of the nearest regions of space has already developed its own history which consists of separately distinguishable stages. During the first stage, mankind's curiosity concentrated mainly on the technical possibilities of overcoming Earth's gravity. During the next stage, the main focus of study centered on the survival of living organisms, including humans, in space using technical devices. The present stage primarily involves mankind's active work during prolonged spaceflights. Hence, in the short history of astronautics and cosmonautics, the centre of interest has been shifting away from the engineering sciences towards the biological, medical, and psychological sciences. [21] (p. 352)

Along with the psychological studies related to stress and workload are studies concerned with manual control and other areas in the general category of human factors. On Space Station Freedom, crewmembers will be remotely operating vehicles, robots, and experiments subject to the peculiarities of zero-g, orbital mechanics, temperature extremes, and hard vacuum. Current research geared toward uncovering and exploring performance aspects of this environment could have large payoffs in the future.

The importance of manual control aspects of spaceflight operations, such as rendezvous and docking, was recognized early in the United States space program. After only three manned flights in the Mercury Program, the Technical Director, Behavioral Sciences Laboratory, Aerospace Medical Research Laboratories, Wright-Patterson AFB concluded "that men can contribute greatly to the successful accomplishment of many types of space missions. . . . the Mercury astronauts were able to manually compensate [sic] for equipment malfunctions and thereby complete missions which otherwise would have failed or terminated prematurely." [22] (p. 79) As Gemini XII and Apollo XI astronaut Buzz Aldrin explains, "Manned orbital rendezvous was a vital field, because any way you cut it, if we were going to assemble large interplanetary spacecraft, we'd have to master the techniques of space rendezvous—bringing two or more separately launched spacecraft together in orbit. With computers we could reduce the blizzard of spherical geometry and calculus equations down to automated rendezvous procedures. But I'd seen enough autopilots malfunction during my flying career to realize that the spacecraft NASA planned to use for Earth orbital lunar spaceflight would need some kind of manual backup." [23]

The Soviets also value the flexibility that manual control allows in "the capabilities of man to see three dimensions and to evaluate the situation better than a machine for flight conditions that have not been provided for by the program." [15] (p. 804) Gemini X and Apollo XI astronaut Michael Collins advocates manual control as follows, "was this not a noble cause, to build an autonomous capability, to allow a manned spacecraft to roam free of ground control, to compute its own maneuvers? Was not the very name of the game, in manned space flight, to put the pilots in control?" [24] (p. 169) Further justification for manual control may be found in the airline industry where "pilots still manually fly even the most highly automated aircraft, if only to maintain their flying skills in the case that they are called on if the automatics fail." [1] (pp.

293-4)

While automation is and will continue to be an important aspect of manned space flight,
> It is unlikely that the pilot will be eliminated, any more than will the operator of a nuclear power plant. Our society believes that humans should have ultimate responsibility for control of complex systems even if inserting the human degrades overall system performance most of the time. The human is still the ultimate back-up system. While machines that are overloaded fail abruptly, people degrade gracefully under excessive levels of workload. Thus it seems prudent to include human operators, even if only as the sub-system of last resort that can "pull the plug." Furthermore, there are also strong political forces to keep humans employed. [25] (p. 183)

The development
> of rendezvous and docking procedures arose in the evolution of the U.S. space program once the initial exploratory phase (Mercury) had been successfully completed and missions became more ambitious. In the United States, the Gemini program was used to acquire these techniques and develop these technologies, and to give astronauts the practice they needed to get to the moon. Orbital rendezvous procedures were performed as the various Gemini craft tracked and approached their respective rendezvous targets. Gemini demonstrated that "precise flight-crew responses during orbital flight is [sic] critically dependent upon the fidelity of the simulation training received prior to flight" [25] (p. 1) [3] (p. 11).

While simulators were, and are, used extensively for procedure development and training, evidence of their use in human factors studies is virtually non-existent in the literature. Presumably, with the rush to get men to the moon and back before the end of the decade, time, money, and effort spent on such studies was not justifiable. With only 12 operational dockings in the Apollo program, three in Skylab, and one for the Apollo-Soyuz Test Project (ASTP), productivity benefits would not have been realized with the low economies of scale of 16 dockings over 10 years. By the end of the 1990s, space shuttle orbiters will be bringing crew and equipment to Space Station Freedom. By 2010, components of lunar bases and spacecraft will be brought to the station for checkout and assembly. Sorties will be made to investigate and repair satellites and other payloads. A small investment now geared toward a better understanding of manual control aspects of piloting docking maneuvers and other space operations could yield large payoffs in the future in terms of safety, reliability, productivity, and launch costs.

Another justification for the current interest in manual control aspects of spacecraft docking operations concerns the differences in on-board sensors and instrumentation between previous Apollo missions and future spacecraft dockings to the station. As Buzz Aldrin recalls the docking after ascent from

461

the lunar surface, "Our radar and the computers on the two spacecraft searched for each other and then locked on and communicated in a soundless digital exchange." [23] Current plans for Space Station Freedom omit this communication capability as well as a laser rangefinder that was proposed earlier. The rendezvous radar on the space shuttle is limited by a minimum operational range of 80 feet. [27] Ironically, rendezvous and docking operations will be harder to perform twenty years from now than they were twenty years ago because of reduced instrumentation and a paucity of information presented to the pilots.

In addition to discovering approach velocities, braking gates, control modes and other flight procedures that will increase safety, efficiency, and productivity, and decrease fuel use, research into the manual control aspects of space operations such as docking maneuvers has hardware implications. For example, there is a tradeoff between the mass of a space station or satellite docking fixture and the amount (mass) of fuel that will be consumed by a vehicle docking with it. (Although current planning has vehicles berthing with the station via a manipulator arm, rather than docking, this tradeoff may be appropriate for satellite dockings.) Increased strength is paid for with an increase in docking fixture mass. More fuel is needed to control impact velocity when docking with a delicate structure than with a more massive, stronger one. Since launch costs are directly proportional to launch mass, hardware designers are incessantly endeavoring to reduce mass. However, over an operational lifetime, operational costs may be elevated as a result of the increased fuel consumption necessary to dock with a lighter, more fragile target. Human factors studies can produce data concerning the fuel mass/approach velocity tradeoff. Flight simulator experiments can be conducted to analyze quantitatively the effect that impact velocity has on fuel consumption. In this way, the lifetime operational costs can be better understood and long-term benefits will not be sacrificed for short-term gains.

From the Gemini program, there is a historic example of uncertainties in fuel consumption requirements. Ratios of actual fuel consumption to theoretically minimum fuel consumption values varied from 1.52 to 4.28 for the ten rendezvous operations. [24] Clearly, mission planners need to have a better idea of this ratio in order to allocate supplies for any given mission correctly. Research into the manual control aspects of rendezvous maneuvers will help reduce both the absolute value and the variance of the actual/theoretical fuel consumption ratio.

Additionally, a comprehensive study of the impact velocity effect on fuel consumption will also yield the effect on mission duration. In the future, the desire to dock during orbital daylight, an increase in space traffic, and other constraints will make time management almost as important as fuel management. In January 1990, the Long Duration Exposure Facility (LDEF) was within weeks of tumbling out of control and deorbiting when the space shuttle crew rescued it. This is one example where time may be very important and a full understanding of the performance envelope for piloting may be necessary for the success of the mission. Studies can be performed to assess the impact of docking port location, number, and design on time and fuel consumption. In short, a comprehensive and extensive study of manual control aspects of spaceflight can produce many long-term savings of time, fuel, and launch costs while increasing safety and reliability. This is currently a timely research agenda to which greater resources and attention are owed.

BACKGROUND

The first spacecraft docking occurred in March 1966 during the Gemini 8 flight of Dave Scott and Neil Armstrong. Armstrong piloted the docking to the Agena target vehicle. "It was also 100 percent manually flown, not unlike mid-air refueling of airplanes, and it made us pilots feel good to hear Neil report that it had been easy, with no surprises." [24] (p. 180) One half hour after docking, however, a malfunction in the Gemini attitude control system led to uncontrolled tumbling. Armstrong was able to null the motion until he released the hand controller at which point the tumbling restarted. To simplify the problem, he backed the Gemini away from the Agena. This unfortunately aggravated the situation and the rotation rate increased to 300 degrees per second. He was ultimately able to recover control and stop the tumbling solely through manual control of the reentry attitude system. "The whole thing had lasted perhaps ten minutes, but they were the hairiest ten minutes in the space program so far." [24] (p. 182).

"Neil was far and away the most experienced test pilot among the astronauts." [24] (p. 317) His "Right Stuff" piloting skill was also required during the Apollo XI landing when he discovered that the designated landing location was too rough to achieve a safe landing. He then resorted to manual piloting to traverse the craters to a smoother spot. As Armstrong's crewmate Buzz Aldrin recalls,

> At 500 feet, Neil was not satisfied with the landing zone. He took over manual control from the computer, slowing our descent from 20 feet per second to only nine, and then at 300 feet, to a descent of only three and a half feet per second. . . . Neil did not like what he saw below. [23]

Ultimately, of course, the landing was successful illustrating the flexibility of manual back-up without which, the mission most likely would have failed. As Gordo Cooper said after his Faith 7 debriefing, "'. . . man is a pretty good backup system . . .'" [23]. Along with John Glenn's piloting skill in flying the reentry of his Friendship 7 mission when it was thought his heat shield became dislodged, this incident helped to entrench the importance of manual control in the NASA mindset.

Even in the commercial airline industry, where there is far more collective piloting experience than in space, there is an apprehension of automation. Pilots have been known to make comments such as, "In some cases the forces driving

technology have caused the design of automated systems which compromise the ability of the pilot to fulfill his responsibilities for the safety of the airplane under his command." [28] (p. 155) Since all NASA pilots come from a jet pilot heritage, comments such as these are relevant for the space program as well.

On Gemini X, in July 1966, John Young "finds [docking] as easy as Neil did on Gemini 8." [24] (p. 211). Dockings were also performed on Geminis XI, and XII in September and November respectively. Apollos 9 and 10 practiced orbital docking operations with the Apollo configuration in 1969. (See [3] for a detailed description of Apollo rendezvous and docking procedures.)

Despite the flexibility and resourcefulness that crewmembers provide, it must be admitted that they also supply additional means for malfunction and error.

CURRENT STUDY

Docking maneuvers have traditionally been simulated and ultimately performed in a "pulse" control mode. That is, thrusts of a prescribed magnitude (duration) were commanded by deflection of a hand controller regardless of deflection angle or duration. Subsequent burns were only possible after release of the joystick to its rest position. NASA space shuttle pilots and orbital maneuvering vehicle (OMV) pilots currently are instructed to use pulse control presumably for fuel consumption and safety reasons.[27, 29]

Nevertheless, all previous experimentation by the authors involved acceleration control in which thruster commands were sent for the duration of the deflection.[3-8] This study involved a formal comparison between pulse control and acceleration control to determine which is better for fuel consumption, mission duration, safety, and other considerations.

In the current study, the trials were organized in an APPA and PAAP orders where A denotes a series of 18 simulated dockings using acceleration control and P corresponds to a series with pulse mode. Subjects who began with acceleration mode, continued with two blocks of pulse mode before returning to their final block with pulse mode (i.e., APPA). Subjects beginning with pulse mode did the opposite (i.e., PAAP).

One of the intents of this format was to unearth any asymmetrical transfer that may be present. Asymmetrical transfer would be evident if a control mode x order (mode x first mode) effect were found.[30] It specifically means the effect of practice with one control mode on subsequent performance with the other control mode is different for the two possible sequences of activity (i.e., a PA sequence vs. an AP sequence). This could occur, for example, if subjects who began with pulse mode achieved lower mission duration values when they later flew in acceleration mode than those

who began with acceleration mode and followed with pulse mode. Such a finding would be useful for identifying which control mode to use for training as opposed to flight. Additionally, a control mode x range interaction would indicate which mode were better depending upon initial range of the mission. Preliminary data indicated that learning might be easier in pulse mode but better performance characteristics are achieved with acceleration control. Asymmetric transfer effects can also cloud comparison of control modes since the subjects' asymptotic performance may not be accurately reflected by the experimental data.

METHOD

Nine commercial airline pilots served as paid test subjects in this study. Pilots were used because of the expectation that the manual control, attention, discipline, and intelligence skills typically associated with flying would enable them to be superior subjects. In purely subjective terms, however, they performed no better than any other previous group of simulated spacecraft pilots. For example, neither learning nor performance was consistently better than previous groups of subjects.

The study was performed in the Space Station Proximity Operations Simulator at NASA Ames Research Center. This facility simulated a proximity operations control room on a space station. A PDP 11/60 computer in conjunction with an Evans and Sutherland PS II picture system drove three windows. These windows displayed a simulated view out the -V-bar (negative velocity vector) of a space station in a 270 nautical mile circular orbit around the Earth. An accurate star field was visible with representatives down to the fifth magnitude.

A three-degree-of-freedom displacement hand controller was used to command thruster firings on an orbital maneuvering vehicle (OMV) remotely. Buttons on the hand controller were used to control the thruster characteristics for each coordinate axis independently. Thruster values were toggled among 1.0, 0.1, and 0.01 m/s. The subjects used a joystick-mounted trigger to begin each trial.

A head-up display (HUD) containing flight data was superimposed on the center window. Mission duration, velocity increment, 3-axis range and rate, slant range and rate, and thruster values were presented to the subjects.[18-20]

Test subjects performed simulated docking maneuvers of an OMV to a space station from three different ranges on the -V-bar. Each subject used both control modes in blocks of 18 consisting of 3 ranges (50, 100, and 150 m) x 6 repetitions in a latin squares configuration. Five subjects began with acceleration control and 4 began with pulse control. A test session consisted of two blocks with each control mode. The blocks were arranged in an APPA or PAAP order. This yielded a total of 72 trials for each subject. Experimentation required about five hours per subject.

463

RESULTS

Five-way mixed analyses of variance (ANOVA) with one between factor, first mode, and four within factors, mode, block, range, and trial were performed on the data. All statistically significant effects at the .05 level for the complete data set are summarized in the following table. Trial refers to consecutive presentations of identical experimental treatments. Mode, range, and block are the same for a group of six trials. Block distinguishes between both groups of 18 consecutive trials with the same control mode. The blocks were designated first half and second half.

Table I: Significant effects from ANOVA.

Dep. Var.	Significant Factor(s)	F	p
Mission	Mode	12.544	.0094
Duration	Range	24.156	.0001
	Trial	4.143	.0046
	Mode * Block * 1st	5.835	.0464
Velocity	Mode	6.431	.0389
increment	Range	34.57	.0001
	Block * range	5.792	.0147
	Mode * bl * r * tr * 1st	2.100	.0357
X Velocity	Block	7.118	.0321
increment	Range	31.344	.0001
	Trial	2.653	.0390
	Block x range	5.864	.0141
Y Velocity	First Mode	31.523	.0008
increment	Range	6.861	.0084
	Range * First Mode	6.721	.0090
	Mode * Range * Trial	2.308	.0208
	Mode * r * tr * 1st	2.287	.0219
	Mode * bl * r * Trial	1.984	.0481
	Mode * bl * r * t * 1st	2.018	.0441
Z Velocity	Range	4.142	.0429
increment			
X Rate	Trial	2.759	.0334

Control mode produced statistically significant, but opposite, effects on mission duration and Δv. Mission duration was lower with pulse mode while Δv was lower with acceleration mode. Subjects used more fuel to travel faster with pulse mode than with acceleration mode. As in more mundane, Earthbound, linear environments, greater velocities, leading to reduced mission durations, are paid for with increased fuel consumption. Although the subjects were trained to criterion, further training could most likely be used to reduce mission duration and/or fuel consumption levels. These results give some indication of what the underlying tendencies are before extensive training.

Mission duration, velocity increment, X velocity increment, Y velocity increment, and Z velocity increment all increased with range. Subjects commanded the OMV to "fly" at faster rates from further distances. These higher average velocities were paid for by increased fuel consumption.

Z velocity increment, the cumulative total of thrusts used to correct for orbital mechanics effects, increased with initial range. This increase was due to the increase in mission duration with range. More fuel was required to compensate for the orbital mechanics effects when more time was given for them to operate.

The most unusual range effect was the one reflected in Y Δv. The y-axis was the out-of-plane component. Since motion along this axis is uncoupled from motion along the other two axes, an object with zero y displacement with respect to a target needs no attention. Although the trials in this study were initialized so that no thrusts along the y-axis were required, accidental commands were made from which recoveries had to be made to achieve a successful docking. Most likely, the longer mission durations associated with the greater initial ranges provided the subjects with more time in which to cause a y disturbance.

Although the subjects practiced to criterion before experimentation, a practice effect in which subjects improve with experience was still evidenced in the data. Mission duration decreased with trial in a typical learning curve format. Surprisingly, X velocity increment increased with experience. This effect was most likely due to subjects becoming more comfortable with the simulated docking maneuver and consequently using more fuel to travel faster.

The X velocity increment data demonstrated a block effect also. Fuel consumption along the x-axis was less in the beginning of testing than in the end. Values from the first eighteen trials with a mode were less than those from the second half with means (SDs) of 7.7 (4.9), and 8.9 (6.1) m/s for the first half and second half respectively. This effect was similar to the trial effect with fuel consumption and velocity increasing with experience. It shows the trend following experience not only within blocks as with the trial effect but also between blocks as mentioned here.

Three 2-way interactions, two 3-way interactions, two 4-way interactions, and two 5-way interactions also resulted from the data analysis. Higher order effects are typically difficult to decipher. Of particular interest are the ones containing a mode or first mode term.

The mode x block x first mode interaction for the mission duration data appears in Figure 1. It shows that the main effect relationship between the modes, namely, the mean for mission duration in pulse mode is less than the acceleration mean only holds for the first half of the data. In the second half of the data, the pulse data for subjects who began in acceleration mode has the same mean as the data from those subjects who began in pulse mode.

The error bars indicate that the data for blocks 2 and 3 for both sets of subjects are not distinct. Essentially, mission duration values for the middle two blocks are the same for both modes. There is also no statistically significant distinction between the data from blocks 1 and 4 in the PAAP group. However, the

I

mission duration mean for block 4 is lower than that for block 1 in the APPA group. (See Figure 1.)

First Mode x Block x Mode Interaction
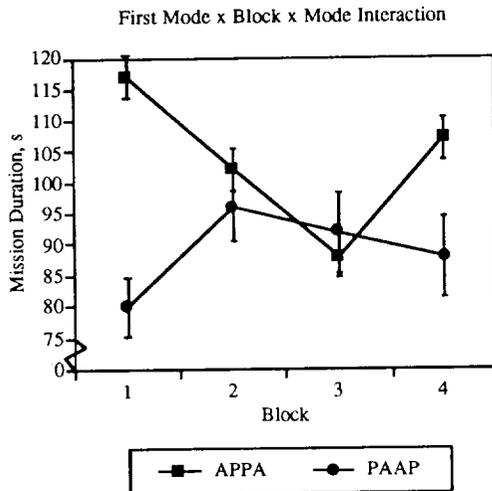
Figure 1: Mission duration 3-way interaction illustrating asymmetrical transfer.

No improvement in mission duration was found for the subjects who began with pulse mode while the data from the acceleration first group display learning. These data support the experimental hypothesis that experience in pulse mode helps performance in acceleration mode. Asymmetrical transfer was found in that the mode transitions could not be predicted solely from the main effect. Both acceleration means in the PAAP group were lower than both acceleration means for the APPA group. The last pulse mission duration means for both subject groups were equal (i.e., block 3 in the APPA group and block 4 in the PAAP group).

Analysis of the transitions between consecutive blocks also yielded interesting results. Both PA transitions were of the same positive slope. While this is illustrative of the main effect, (that is, pulse mission duration lower than acceleration mission duration), only one of the AP transitions was significantly downward. The single PP transition was downward, again indicating a learning benefit from a previous experience with pulse. Conversely, the single AA transition was unchanged.

An ANOVA was performed on the data collapsed across block and trial to determine which combinations of independent variables were more likely to cause an unsuccessful mission. No statistically significant effects were uncovered. Neither mode was found to be inherently safer than the other. No combination of range and mode was more conducive to errors than any other.

The finding that fuel consumption levels, measured as velocity increment or Δv, were lower in acceleration mode than in pulse mode corroborates the results from the preliminary experimentation. Pulse mode is not inherently more fuel conservative than acceleration mode as one might presume from studying the appropriate NASA manuals.[1, 2] This indicates that fuel can be used more efficiently in acceleration mode than pulse mode in a docking operation. This is probably due to the greater dynamic range with acceleration control allowing for greater flexibility and fine tuning capability.

The asymmetrical transfer discovered here is important for researchers investigating the impact of control modes on spacecraft docking operations. This result should be regarded as a forewarning that investigators should be careful when designing experiments and formulating conclusions. The asymmetry illustrates an inconsistent main effect for which one must account before attributing a result to a control mode. In comparing different control modes, experimenters should be sure to provide sufficient intervening practice to prevent the effects of asymmetrical transfer from contaminating the experimental results.

The data from this study demonstrated that dockings could be performed faster, albeit at the expense of greater amounts of fuel, in pulse mode than in acceleration mode. While the absolute values of time and fuel were specific to the thruster values that were used, this relationship should be preserved with different thrusters. A whole assortment of studies could be performed to examine the effect that thrusters with different magnitudes from the ones simulated here have on the data. An interaction between thruster size and range might also be revealed. What is clear, however, is that pulse mode is not definitively more fuel efficient than acceleration mode in all situations. Probably the most necessary conclusion to be made at this point is the requirement of further human factors and manual control experimentation before flight protocols are generalized for all vehicles in all situations.

REFERENCES

1.   Nagel. D. C., "Human error in aviation operations," In E. L. Wiener & D. C. Nagel (Eds.), HUMAN FACTORS IN AVIATION, Academic Press, Inc, San Diego, CA, 1988, pp. 263-303.

2.   Birch, S. "Human factors in cockpit design," AEROSPACE ENGINEERING, May 1990, pp. 45-49.

3.   Brody, Adam, "Spacecraft Flight Simulation: A Human Factors Investigation into the Man-Machine Interface Between an Astronaut and a Spacecraft Performing Docking Maneuvers and Other Proximity Operations," Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, MA., April 1987; see also NASA-CR-177502, Sept. 1988.

465

4. Brody, Adam, "The Effect of Initial Velocity on Manually Controlled Remote Docking of an Orbital Maneuvering Vehicle to a Space Station," AIAA Paper 89-0400, January 1989.

5. Brody, Adam, "Remote operation of an orbital maneuvering vehicle in simulated docking maneuvers," NASA CP-3059, 1989, pp. 471-475.

6. Brody, Adam, "Evaluation of the "0.1% rule" for Docking Maneuvers," JOURNAL OF SPACECRAFT AND ROCKETS, Vol. 27, No. 1, 1990, pp. 7-8.

7. Brody, Adam & Ellis, Stephen, "Manual Control Aspects of Space Station Docking Maneuvers," SAE Technical Paper 901202, July 1990.

8. Brody, Adam & Ellis, Stephen, "Effect of an Anomalous Thruster Input During a Simulated Docking Maneuver," JOURNAL OF SPACECRAFT AND ROCKETS, Vol. 27, November/December, 1990, pp. 630-633.

9. Brody, A. R., "EivaN: An interactive orbital trajectory planning tool.," JOURNAL OF SPACECRAFT AND ROCKETS, Vol. 27, November/December, 1990, pp. 681-683.

10. Clark, H. J., "Space rendezvous using visual cues only," (Aerospace Medical Research Laboratory TR-65-10), 1965.

11. Ellis, S. R. and Grunwald, A. J., "The dynamics of orbital maneuvering: design and evaluation of a visual display aid for human controllers, PROCEEDINGS OF THE AGARD FMP SYMPOSIUM ON SPACE VEHICLE FLIGHT MECHANICS. November 13-16, 1989 Luxembourg, 29-1 - 29-13.

12. Grunwald, A. J. & Ellis. S. R., "Interactive orbital proximity operations planning system," (NASA-TP-2839). National Aeronautics and Space Administration, Washington, DC, 1988.

13. Haines, R. F., "Space vehicle approach velocity judgments under simulated visual space conditions," (Technical Memorandum 89437). National Aeronautics and Space Administration, Washington, DC, 1987

14. Hartley, C. S., Cwynar, D. J., Garcia, K. D., Schein, R. A., "Capture of satellites having rotational motion," PROCEEDINGS OF THE HUMAN FACTORS SOCIETY 30TH ANNUAL MEETING, The Human Factors Society, Santa Monica, CA, Inc., 1986, (pp. 875-879).

15. Meshcheryakov, I. P., & Minaev. S. A., "Construction and investigation of an information model of the process of approach of piloted spacecraft," KOSMICHESKIE ISSLEDOVANIYA, 15, 1976, pp. 804-807.

16. Meshcheryakov, I. P., Sal'nitskii, V. P. & Nechaev, A. P. Engineering-psychological study of information imaging systems. KOSMICHESKIE ISSLEDOVANIYA, 16, 1978 pp. 453-455.

17. Novikov, U., "Docking in space a complex problem," Department of Commerce: Joint Publications Research Service, 1968.

18. Brody, Adam, "Modifications to the NASA Ames Space Station Proximity Operations (PROX OPS) Simulator," NASA-CR-177510, October 1988.

19. Haines, Richard, "Design and Development of a Space Station Proximity Operations Research and Development Mockup," SAE Technical Paper 861785, October 1986.

20. Lee, Ed, & Wu, Amy, "Space Station Proximity Operations Workstation Docking Simulation," Sterling Software, TN-87-7104-519-13, Palo Alto, CA, March 1987.

21. Terelak, J. F., "Trends in Poland in space psychology research," AVIATION SPACE AND ENVIRONMENTAL MEDICINE. 60, April 1989, pp. 352-360.

22. Grether, W. F., "Human performance capabilities for military operations in space," Lectures in Aerospace Medicine, USAF School of Aerospace Medicine, Aerospace Medical Division, Brooks Air Force Base, TX., 1963, pp. 77-102.

23. Aldrin, B. & McConnell, M., MEN FROM EARTH, Bantam Books, New York, NY, 1989 .

24. Collins, Michael, CARRYING THE FIRE AN ASTRONAUT'S JOURNEYS, Bantam Books, New York, NY, 1974.

25. Kantowitz, B. H. & Casper, P. A.,"Human workload in aviation," In E. L. Wiener & D. C. Nagel (Eds.), HUMAN FACTORS IN AVIATION, Academic Press, Inc, San Diego. CA, 1988, (pp. 157-187).

26. Mueller, G. E., "Introduction," GEMINI SUMMARY CONFERENCE, NASA SP-138, Manned Spacecraft Center, Houston, TX., February 1-2, 1967, pp. 1-3.

27 Oberg, James, "Rendezvous and Proximity Operations Handbook," NASA Lyndon B. Johnson Space Center Mission Operations Directorate Flight Design & Dynamics Division JSC-10589, May 1988.

28. Hoagland, M., "Winging it in the 1980s: Why guidelines are needed for cockpit automation," THIRD AEROSPACE TECHNOLOGY CONFERENCE PROCEEDINGS, Warrendale, PA, 1984, pp. 155-162

29 Sedej, Daniel & Clarke, Steven, "Rendezvous/Proximity Operations Workbook RNDZ 2102," NASA Lyndon B. Johnson Space Center Mission Operations Directorate Training Division Flight Training Branch, 1985.

30 Poulton, E. C., Tracking Skill and Manual Control, Academic Press, Inc., New York, 1974.

466

I

Session H2: SPACE PERCEPTION AND PHYSIOLOGY - I

Session Chair: Dr. Samuel Schiflett

EFFECT OF MICROGRAVITY ON SEVERAL VISUAL FUNCTIONS
DURING STS SHUTTLE MISSIONS

Melvin R. O'Neal, OD, PhD, Lt Col, USAF
H. Lee Task, PhD
Louis V. Genco, OD, MS, Col, USAF

Optical Systems Branch, Human Engineering Division
Crew Systems Directorate, Armstrong Laboratory
AL/CFHO, Wright-Patterson AFB, Ohio 45433

ABSTRACT

Many astronauts and cosmonauts have commented on apparent changes in their vision while on-orbit. Comments have included statements of supposed improved distance acuity to decreased near vision capability. The purpose of this study was to assess not only changes in visual acuity, but expand the assessment to several other visual functions for a comprehensive battery of tests. Vision was assessed using an innovative device, the Visual Function Tester - Model 1 (VFT-1), which presents the tests at optical infinity and includes critical flicker fusion, stereopsis to 10 seconds-of-arc, visual acuity in small steps to 20/7.7, cyclophoria, lateral and vertical phoria, and retinal rivalry. Vision was assessed 2 times prelaunch at L-14 days and L-7 days, 3-4 times while on-orbit, at landing, and 2 times postlanding at L+3 days and L+7 days. There were 26 STS astronauts that participated, with data on 20 astronauts used for analysis. There was a typical wide variability between subjects in baseline visual performance for each parameter at the prelaunch sessions. There was a slight but statistically significant decrease in visual acuity while on-orbit that was not clinically significant. For stereopsis (ie. depth perception), there was a small improvement on-orbit that was not statistically significant. There were no changes during space flight for any of the other visual parameters tested. A few individuals showed apparent changes in acuity and stereopsis. The possibility exists that microgravity affects the visual system of some individuals differently, as with space adaptation syndrome. Repeat data on 2 astronauts showed good repeatability between the 2 flights. These results pertain to only short term space flight on the STS shuttle, and longer flights are necessary to determine if there is any relationship between mission duration and these visual functions.

INTRODUCTION

Many of the astronauts and cosmonauts have commented on apparent changes in their vision while on-orbit. Some comments by some astronauts[1] and cosmonauts[2] have included descriptions of earth features and objects that would suggest enhanced distance visual acuity. On the other hand, some cosmonaut observations suggest a slight loss in their visual discrimination during initial space flight[3]. In addition to distance vision changes, astronauts have mentioned a decreased near vision capability that either improved during spaceflight or did not recover to normal until return to earth [4].

In the late 1960's, Duntley et al.[5] used both ground targets and a hand-held device to assess visual acuity during Gemini V and VII and concluded there was no significant change in acuity. The Duntley device presented high and low contrast bar targets at optical infinity. The Soviets have tested acuity during a number of Voskhod and Soyuz flights, with the results summarized by Lazarev[6]: (1) slight decrease of 5 - 10% in high contrast acuity for 2 subjects on

Voskhod flights, (2) 10% reduction in acuity for both high and low contrast targets for 3 subjects, but 20% increase in high contrast acuity for one subject on Soyuz-4 and 5, and (3) 18% drop in high contrast acuity and only 4% drop in low contrast acuity for one subject on Soyuz-9. However, the Soviet tests were conducted at a distance of about 30 cm, and therefore are for near vision. Since many cosmonauts are older in age, their findings may be complicated by any near vision problems.

These tests were performed at different optical distances and the results may have been affected by other factors, such as age and lighting conditions; and thus, are not comparable. The purpose of this study was to not only assess the effect of microgravity on distance visual acuity, but expand the assessment to several other visual functions under controlled conditions and over the length of space flight. The hand-held device used presented all targets at optical infinity and set light levels, and included high contrast acuity targets in small size increments, flicker fusion, stereopsis, phorias, and retinal rivalry, for a comprehensive battery of tests.

METHODS

Subjects

To date, 26 STS astronauts have participated since 1984 during 7 missions; and the mission number, number of subjects, and flight days on which data taken are listed in Table I. Unfortunately, 5 subjects from the same mission were able to complete only one pre-flight test close to launch and only one on-orbit test, and are not included in the analysis. Also one subject experiencing toric soft contact lens problems was eliminated, giving a total of 20 subjects for analysis. Visual correction for distance was worn, if necessary, when using the test device. Three astronauts wore contact lenses for all tests, including on-orbit. There was no statistical difference for these contact lens wearers from the other subjects and their data was included in the overall analysis. Two astronauts repeated the device on second missions, allowing at least limited repeat comparison.

Apparatus

Vision was assessed using a new innovative device, the Visual Function Tester - Model 1 (VFT-1) developed in our Laboratory by Drs Task and Genco. The VFT-1 has been previously described in detail[7], and an external schematic is shown in Figure 1. The battery of tests was chosen on the basis of detecting small changes in neurological or muscular balance. The visual parameters tested are presented at optical infinity and include critical flicker fusion, stereopsis (depth perception) to 10 seconds-of-arc, visual acuity in small steps to 20/7.7, cyclophoria, lateral and vertical phoria, and retinal rivalry (ocular dominance). The unit is self-powered (battery), and uses microelectronic circuitry and LED light bar modules to illuminate the test patches within the device.

Procedure

A pre-mission briefing and instrument familiarization session was conducted 1-2 months before the mission. Vision was assessed 2 times pre-flight at 14 days (L-14) and 7 days (L-7) before launch. While on-orbit data taking occurred 3-4 times; after wake-up and daily when possible, but were spread out on some missions. Post-flight, the VFT-1 was performed as part of the medical examination conducted about 2 hours after landing and again 3 days (L+3) and 7 days (L+7) after landing. All pre-flight and post-flight tests were performed with one of the experimenters in attendance.

Data Analysis

The data was analyzed by calculating the difference between the mean of the two pre-flight sessions (taken as baseline) and each subsequent measurement for each individual. Not all individuals performed the tests at the same mission elasped time on-orbit. There was also a slight non-normality due to the nature of some tests and, given the reasonably large number of subjects, nonparametric analysis (Wilcoxon Signed-Rank) was used to test for statistical significance.

Table I.
VFT-1 HISTORY OF SPACE-FLIGHT

MISSION	(DATE)	NO. SUBJECTS	DATA DAYS
STS-41D	(AUG 84)	2	1, 2, 4, 6
STS-41G	(OCT 84)	4	2, 5, 8
STS-51C	(JAN 85)	5	2
STS-51J	(OCT 85)	4	2, 3, 4
STS-33	(NOV 89)	5	1, 2, 3, 4
STS-36	(FEB 90)	5	2, 3, 4, 5
STS-38	(NOV 90)	3	2, 3, 4

Visual Function Tester No. 1

CONFIGURATION
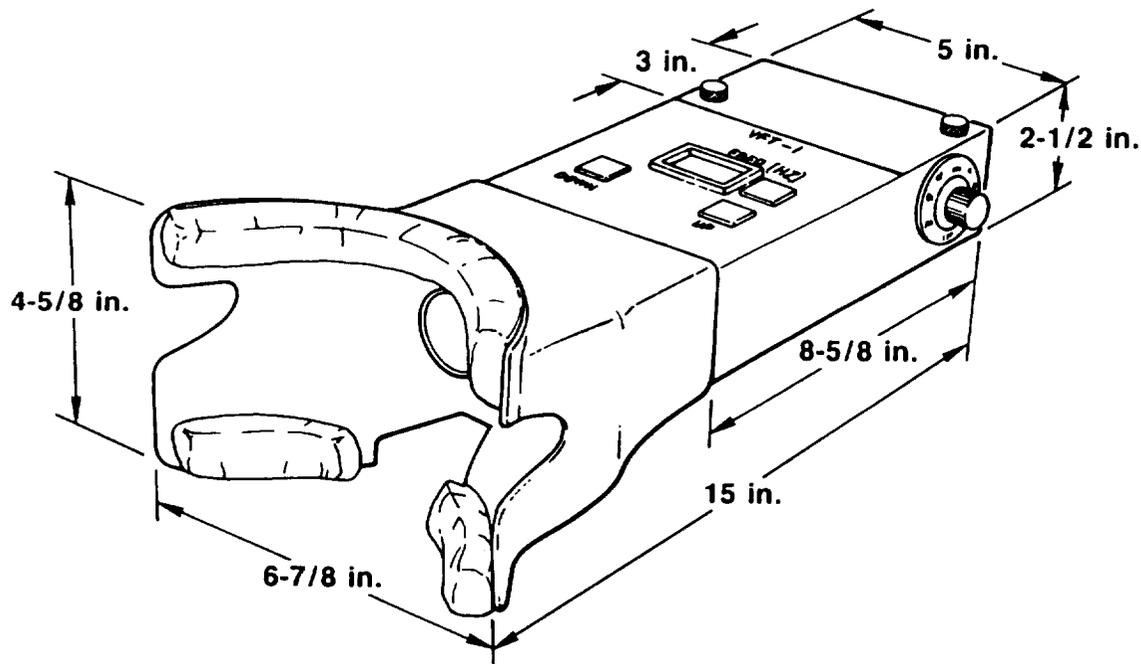
Figure 1. VFT-1 external view.

RESULTS

The group data for visual acuity, stereopsis, lateral and vertical phoria, cyclophoria, and critical flicker fusion is shown in Figure 2. The corresponding data days are: L-14 days is Pre-Flight 1, L-7 days is Pre-Flight 2, on-orbit data is at hours of mission elapsed time (MET), Landing is Post Flight 1, L+3 days is Post Flight 2, and L+7 days is Post Flight 3. The size of the dots represents the number of subjects with the same performance. It should be remembered that the subjects performed the tests a different number of times and at different MET on-orbit. The retinal rivalry data was obtained as the number of pattern reversals in a timed interval for some subjects and by subjective assessment on other subjects, and is not shown.

The group data show the normal wide variability between subjects in baseline visual performance for each parameter at the pre-flight sessions, as is characteristic of psychophysical data. Any apparent trend in the on-orbit data seen in the figures may not be actual, since less subjects performed the tests at the longer MET. The data was analyzed by calculating the difference between the mean of the two pre-flight sessions (taken as baseline) and each subsequent measurement for each subject. As seen in Figure 3, there was no apparent trend in the difference from the pre-flight mean for both on-orbit and post-flight individual data for lateral and vertical phoria, cyclophoria, and critical flicker fusion. Statistical analysis also showed no significant difference. There was also no difference in these comparisons for retinal rivalry.

470
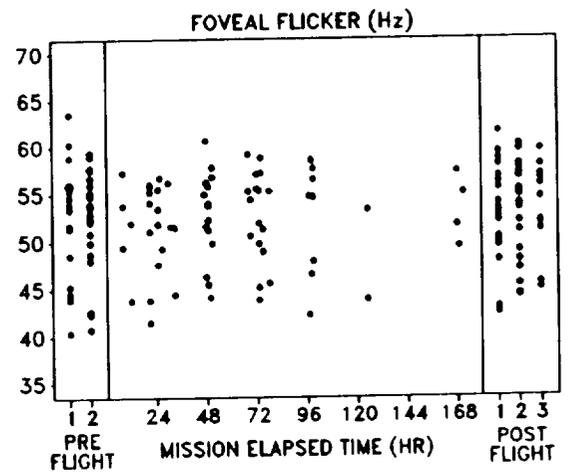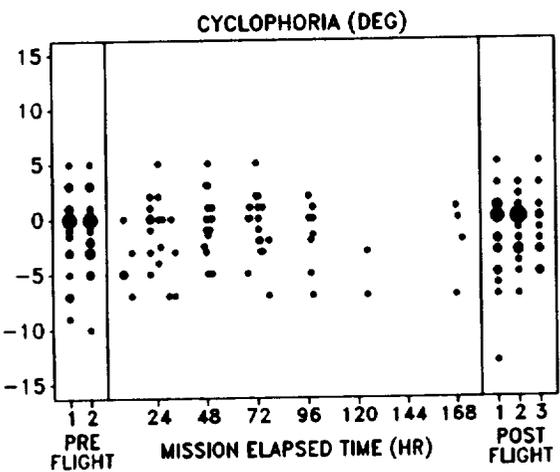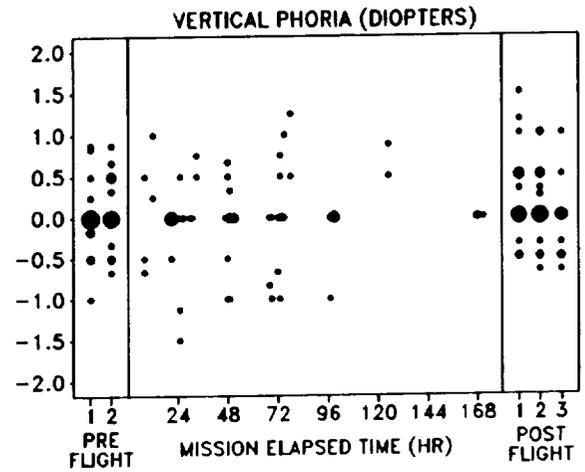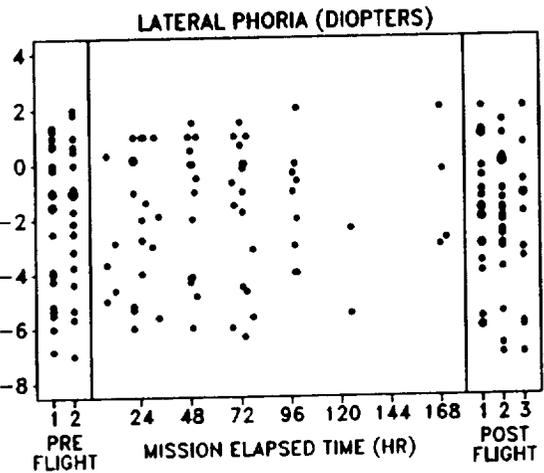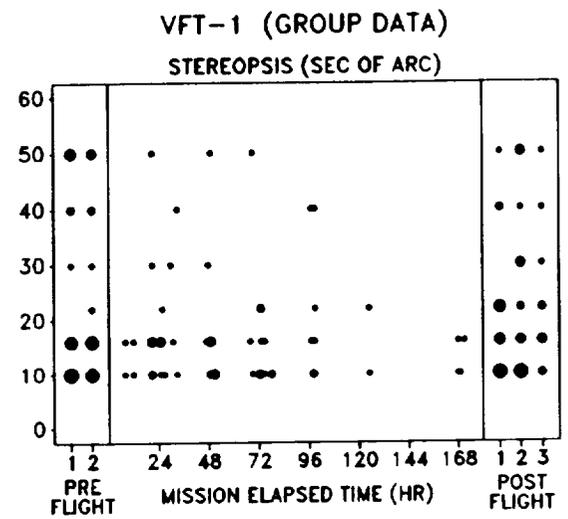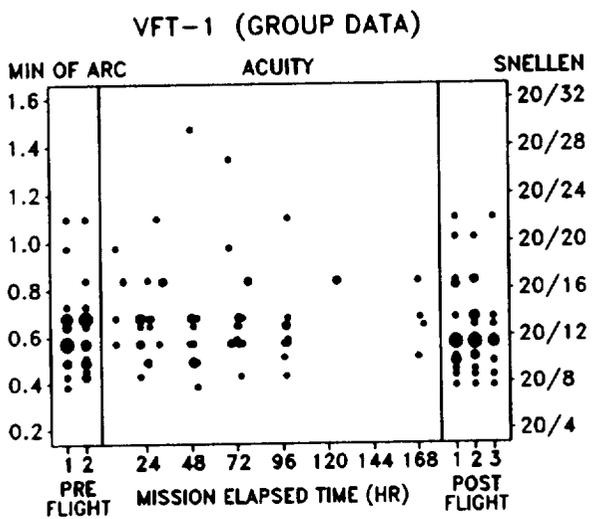
Figure 2. Group data on six vision tests for 20 astronauts. Size of the dots represents the number of subjects with the same performance.

471
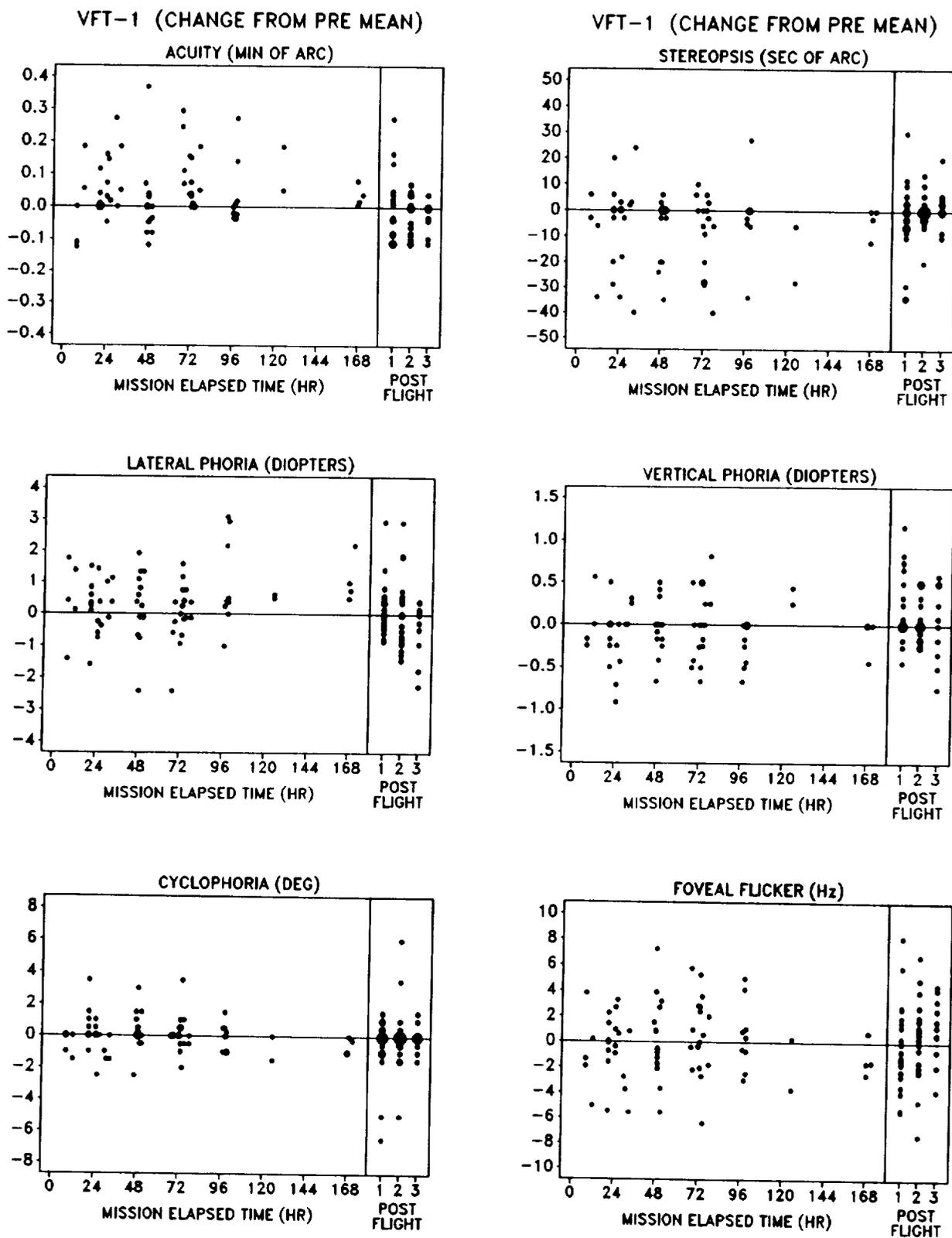
Figure 3. Difference between mean of two pre-flight sessions (baseline) and each subsequent measurement on six vision tests for each of 20 astronauts.

472

The mean pre-flight (baseline) and mean on-orbit change from pre-flight for the six vision tests is given in Table 2. As a group, the visual acuity was excellent at 20/12.5 and the other visual parameters were within normal limits, although slightly poorer for stereopsis.

Stereopsis Change

The difference in stereopsis from the pre-flight mean of the on-orbit and post-flight individual data is also shown in Figure 3. There is a slight trend towards smaller seconds-of-arc stereopsis (i.e. improved stereopsis) while on-orbit, that is no longer appparent post-flight at landing or the later data sessions.

The mean change in stereopsis from the baseline at the subject's first data take on-orbit was -5.0 arc sec and was the same (p=0.99), also at -5.0 arc sec, at the subject's last data take on-orbit. The data was therefore combined to obtain group change on-orbit. The mean group change in stereopsis was -4.9 arc sec from baseline while on-orbit, which was nearly significant (p=0.07). The mean change in stereopsis from baseline was only -0.8 arc sec at landing and +1.1 arc sec by the second post flight (L+3 days) session.

Visual Acuity Change

The difference in visual acuity from the pre-flight mean of the on-orbit and post-flight individual data is also shown in Figure 3. There is a definite trend toward larger minute-of-arc resolution (i.e. decreased acuity) while on-orbit, that is no longer apparent post-flight at landing or the later data sessions.

The mean change in visual acuity from baseline at the subject's first data take on-orbit was +0.04 min arc, but is not significant (p=0.13). At the subject's last data take on-orbit, the mean change in acuity from baseline was +0.07 min arc, and is statistically significant (p=0.001). Comparison of the on-orbit acuity data showed no significant difference (p=0.15) between the first and last data takes, and the data was combined to obtain the group change on-orbit. There was a significant (p=0.005) mean group change in acuity of +0.06 min arc from baseline while on-orbit. However, this corresponds to only a Snellen acuity change of from 20/12.2 at baseline to 20/13.4 on-orbit, which is not operationally significant. There was no change in acuity from baseline at landing (p=0.90) and only slight differences for the other post flight means.

Table II.

VFT-1 GROUP DATA

	MEAN PRE-FLIGHT	MEAN CHANGE
VISUAL ACUITY	0.610 min arc (20/12.2)	+0.06 min arc (to 20/13.4)
STEREOPSIS	19.8 arc sec	-4.9 arc sec
LATERAL PHORIA	-2.08 $^\Delta$ (ESO)	+0.36 $^\Delta$
VERTICAL PHORIA	0.04 $^\Delta$	-0.07 $^\Delta$
CYCLOPHORIA	-1.14 $^\circ$ (ENCYCLO)	-0.02 $^\circ$
FOVEAL FLICKER	52.43 Hz	-0.06 Hz

473

Percent (%) Acuity Change

The percent (%) difference from the pre-flight mean of the on-orbit and post-flight individual data for visual acuity is shown in Figure 4. Most of the data points on-orbit show a positive percent change in min-of-arc letter size, corresponding to a decrease in visual acuity. Single data points on-orbit ranged from a 40% loss in acuity to a 20% improvement in acuity. The mean percent change (loss) in acuity from baseline while on-orbit was 7.5%.

Of particular interest in the repeat subjects was the stereopsis data, since both subjects had shown a marked improvement in stereopsis on-orbit during their first mission. As seen in Figure 5, the pre-flight stereopsis was very similar (although poor) prior to both missions, and although one on-orbit data point for each subject varied, the overall results indicate an improvement in stereopsis occurred again on the second mission.

Figure 4. Percent (%) change in visual acuity from the pre-flight mean for 20 astronauts.

Repeat Subjects

There were two astronauts that participated in the VFT-1 study on second missions, allowing for a repeatability comparison of their data. In general, these subjects confirmed their initial results, as seen in Figure 5 for visual acuity and stereopsis. The filled symbols are for the first mission, while the empty symbols are for the second mission, with circles for one subject and triangles the other. As seen for visual acuity, there were slight differences in baseline visual performance over the years between flights for all parameters tested, with similar change on-orbit.

Figure 5. Individual visual acuity and stereopsis data of two astronauts on first (filled symbols) and second (empty symbols) missions.

474

DISCUSSION

There were no group changes on-orbit for lateral and vertical phorias, cyclophoria, critical flicker fusion, and retinal rivalry. There was a definite trend toward loss of acuity while on-orbit that was statistically significant. However, the mean visual acuity loss was only 0.06 min-of-arc. This corresponds to only a slight change in Snellen acuity from 20/12.2 at baseline to 20/13.4 on orbit. This degree of acuity loss is of little clinical significance, as it represents less than one line of acuity on a standard chart.

The acuity change also equates to a mean loss of 7.5/corresponds well to that reported by the Soviets for their high contrast targets, although their tests appear to have been conducted at a near test distance. There was also a wide range in single data points on-orbit, ranging from a 40% loss to a 20% improvement in acuity, and some subjects varied by as much as 20% on-orbit. This could explain the similar Soviet variability in response between cosmonauts, particularly if the data was taken only one time on-orbit.

Our results do not entirely agree with those of Duntley et al.[7] They found no statistically significant change in visual acuity on-orbit, while we did; while both changes were very small. This may be due to the small increments between letter sizes in our study (e.g. 20/15, 20/14, etc.), allowing for a finer measurement than the typical Snellen acuity line differences.

There was also a slight trend for stereopsis to improve while on-orbit, although mainly for a few subjects and not statistically significant. This improvement was most noticeable for subjects that had large seconds-of-arc stereopsis at pre-flight. In particular, two subjects had pre-flight stereopsis of about 50 arc sec that improved to 20 or 10 arc sec (the smallest stereo target) while on-orbit, and returned back to baseline post flight. For depth perception to improve, then there would have to be less difference between some aspect of the eyes (e.g. refractive error or eye muscle control) such that the eyes coordinated better. However, the data obtained do not allow the mechanism for the improvement to be determined; although the eye muscle control did not appear to change.

There were two astronauts that participated in the VFT-1 study on second missions, allowing for a repeatability comparison of their data. In general, these subjects confirmed their initial results. Of particular interest was the stereopsis data, since both subjects had shown a marked improvement in stereopsis on their first mission (as noted above). Although one on-orbit stereopsis data point for each subject varied, the overall results indicate an improvement in stereopsis occurred again on the second mission. For the other visual parameters, the pre-flight data varied between missions, but the change from baseline on-orbit remained very similar. It would appear that the data obtained with the VFT-1 device has validity, since repeat data overall confirmed earlier findings.

These results are for the short duration spaceflight of STS missions. It would be of interest to evaluate visual performance over longer periods on-orbit. Using a device such the VFT to present test patterns at set light levels and optical distance would allow an accurate determination of the effect of long-term microgravity on a number of visual parameters.

ACKNOWLEDGMENTS

DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense of the U.S. Government.

REFERENCES

1. Zink DL. Visual experiences of the astronauts and cosmonauts. Human Factors, June:187-201, 1963.

2. Lazarev AI and Kovalenok VV. The earth's atmosphere from the Salyut 6. Leningrad, Gidrometeoizdat, 7-30, 1981.

3. Rodikov V. Effects of prolonged space flight on visual perception discussed. Moskovshii Komsomolets, 67:4, 1981.

4. Pool SS and Nicogossian A. Biomedical results of the space shuttle orbital flight test program. Aviat Space and Environ Med, 54:541-549, 1983.

5. Duntley SQ, Austin RW, Harris JL, and Taylor JH. Experiments on visual acuity and the visibility of markings on the ground in long duration earth orbital space flight. NASA CR-1134, 1969.

6. Lazarev AI. Vision in space. Leningrad, Opticheskiye Issledovaniya v Kosmose, 66-87, 1979

7. Genco LV and Task HL. Testing changes in visual function due to orbital environment. Technical Report No. AFAMRL-TR-84-049, 1984. AL/CFHO, Wright-Patterson AFB, Ohio, 45433.

The Neurochemical Basis of Photic Entrainment

of the Circadian Pacemaker

Michael A. Rea, Ph.D
Becky Buckley
Sustained Operations Branch
Armstrong Laboratory
Brooks AFB, TX 78235-5301

Lewis M. Lutton, Ph.D.
Mercyhurst College
Glenwood Hills
Erie, PA 16546

Abstract

Circadian rhythmicity in mammals is controlled by the action of a light-entrainable pacemaker located in the basal hypothalamus, in association with two cell clusters known as the suprachiasmatic nuclei (SCN). In the absence of temporal environmental cues, this pacemaker continues to measure time by an endogenous mechanism (clock), driving biochemical, physiological and behavioral rhythms that reflect the natural period of the pacemaker oscillation. This endogenous period usually differs slightly from 24 hours (i.e., circadian). When mammals are maintained under a 24 hour light-dark (LD) cycle, the pacemaker becomes entrained such that the period of the pacemaker oscillation matches that of the LD cycle. Potentially entraining photic information is conveyed to the SCN via a direct retinal projection, the retinohypothalamic tract (RHT). RHT neurotransmission is thought to be mediated by the release of excitatory amino acids (EAA) in the SCN. In support of this hypothesis, recent experiments using nocturnal rodents have shown that EAA antagonists block the effects of light on pacemaker-driven behavioral rhythms, and attenuate light-induced gene expression in SCN cells. An understanding of the neurochemical basis of the photic entrainment process would facilitate the development of pharmacological strategies for maintaining synchrony among shift workers in environments which provide unreliable or conflicting temporal photic cues, such as the proposed space station.

Photic Entrainment of the Circadian Pacemaker

Considerable evidence suggests that a major light-entrainable circadian pacemaker is located in the ventral hypothalamus in association with the suprachiasmatic nuclei (SCN; Rusak and Zucker, 1979; Meijer and Rietveld, 1989). Bilateral destruction or surgical isolation of the SCN results in the permanent disruption of circadian rhythms in mammals (Inouye et al., 1979; Rusak and Zucker, 1979). Furthermore, transplantation of the SCN from a fetal donor into the hypothalamus of an SCN-lesioned host restores rhythmcity (Sawaki et al., 1986; Lehman et al, 1987; Decousey and Buggy, 1989), and the period of the restored rhythm matches that of the donor (Ralph et al., 1990). In addition, the isolated SCN continues to display circadian behavior in vitro. Circadian rhythms in the neuronal activity (Green and Gillette, 1982; Gillette and Reppert, 1987), neuropeptide release (Earnest and Sladek, 1986), and metabolic activity (Newman and Hospod, 1986) have been demonstrated to persist for several days in cultured SCN explants. These observations strongly suggest that the biological mechanism responsible for the generation of physiological circadian oscillations is an intrinsic component of the mammalian SCN.

Photic entrainment of circadian rhythms occurs as a consequence of the phase specific effects of environmental light on the activity of the circadian pacemaker. This relationship is defined by the phase response curve to light pulses administered to animals maintained under constant darkness (Daan and Pittendrigh, 1976; Takahashi et al, 1984). In nocturnal rodents, light pulses administered during the early subjective night cause phase delays of the pacemaker while pulses delivered during the latter half of the subjective night cause phase advances (Daan and Pittendrigh, 1976; Takahashi et al., 1984). Light pulses given during the middle of the subjective day do not cause phase shifts. Light-induced shifts represent long-term alterations in pacemaker function.

Photic information is conveyed to the SCN through at least two visual pathways. The retinohypothalamic tract (RHT) is a direct, bilateral monosynaptic projection from retinal ganglion cells to neurons in the SCN and the surrounding hypothalamus

(Moore and Lenn, 1971; Johnson et al., 1988a). In addition, a second, indirect visual projection, the geniculohypothalamic tract (GHT) has been described. This pathway projects from the retina to relay neurons in the intergeniculate leaflet (IGL) of the thalamus, which, in turn, send their axons to neurons in the SCN (Swanson et al., 1974; Card and Moore, 1982; Pickard, 1982). Lesion studies have shown that the RHT is both necessary and sufficient to support photic entrainment of circadian rhythms in experimental rodents (Johnson et al., 1988b).

The Neuropharmacology of Photic Entrainment

In addition to light, a number of synthetic and natural neuroactive substances have been tested for their ability to reset the circadian pacemaker. Benzodiazepines (Turek and Losee-Olson, 1986), melatonin (Cassone et al., 1986), theophylline (Ehret et al., 1975), and various neuropeptides (Albers et. al., 1984; Albers et al., 1991) have all been shown to alter the phase of circadian oscillations. However, only a few neurotransmitter-specific agents have been systematically investigated for their effects on light-induced phase alterations.

Gamma-amino butyric acid. A large proportion of the neurons in the SCN contain glutamic acid decarboxylase (van den Pol and Tsujimoto, 1985), the enzyme responsible for the synthesis of the inhibitory amino acid neurotransmitter, gamma-amino butyric acid (GABA). Although it is clear that GABA does not play a direct role in RHT neurotransmission, the abundance of GABA-ergic neurons in the SCN raise the possibility that this neurotransmitter may participate in the processing of photic information in the SCN. In fact, neurophysiological evidence suggests that GABA modulates retinal input to the SCN (Shibata et al., 1986), perhaps by acting presynaptically to regulate neurotransmitter release from RHT terminals (Ralph and Menaker, 1989). For this reason, Ralph and coworkers (1985; 1986; 1989) have investigated the effects of specific GABA agonists and antagonists on light-induced phase alterations of the free running activity rhythm in rodents. These investigators reported that (a) GABA A antagonists attenuate light-induced phase delays (Ralph and Menaker, 1985; 1989), (b) benzodiazepines attenuate light-induced phase advances (Ralph and Menaker, 1986; 1989), and (c) GABA B agonists block both light-induced phase advances and delays (Ralph and Menaker, 1989). These results strongly support a role for the SCN GABA-ergic system in modulation of photic input to the circadian pacemaker.

Acetylcholine. Initial investigations into the identity of the RHT neurotransmitter focused on acetylcholine (ACh). Cholinergic agonists have been reported to mimic (Zatz and Herkenham, 1981; Earnest and Turek, 1983), and antagonists to block (Zatz and and Brownstein, 1981; Keefe et al., 1987), the effects of light on the circadian system. Furthermore, Earnest and others (1985) reported that the phase response curve for intraventricular injections of the cholinergic agonist, carbachol, are similar to the phase response curve for light pulses. This similarity was offered as evidence that ACh might be the RHT neurotransmitter. However, neither retinal ganglion cells nor the optic nerves contain measurable quantities of choline acetyltransferase (Hebb and Silver, 1956), the synthetic enzyme for ACh, and bilateral lesions of the optic nerves do not alter the levels of cholinergic markers in the rat SCN (Rea, unpublished observations). On the other hand, the SCN do receive a cholinergic projection, possibly from the basal forebrain (Ichikawa and Hirata, 1986). One interesting possibility is that ACh from another source may modulate RHT neurotransmission by acting presynaptically (Rusak and Bina, 1990).

Excitatory Amino Acids. Excitatory amino acids (EAA) remain the best candidates for the RHT neurotransmitter. EAA antagonists block fast EPSPs (Kim et al., 1989) and field potential responses (Cahill and Menaker, 1987) of SCN neurons to optic nerve stimulation in the hypothalamic slice preparation. Furthermore, using the same preparation, Liou et al. (1986) have reported that optic nerve stimulation causes the release of radioactivity from SCN slices preloaded with radiolabeled EAAs. Recently, Colwell and colleagues (1990) reported that intraperitoneal injections of the non-competitive EAA antagonist MK-801 attenuated light-induced phase shifts of the free-running activity rhythm in the mouse. In order to determine whether EAA antagonists inhibit light-induced phase shifts by acting on the SCN, we determined the effect of direct injections of EAA antagonists into the SCN on light-induced phases advances of the free-running activity rhythm in hamsters.

Microinjection of EAA Antagonists into the SCN Attenuate Light-Induced Phase Advances

Syrian hamsters were implanted with 26 gauge guide cannulae stereotaxically aimed at the SCN. The cannulae were fixed in place with dental cement and the animals were allowed to recover for 7 –

10 days under LD 14:10. After the
recovery period, the hamsters were
transferred to individual cages equipped
with computer-monitored running wheels
and maintained under constant darkness
(DD). Only animals with robust free
running activity rhythms and stable
periods were used in the study. Hamsters
remained in DD for 7 - 10 days before
treatment. At mid subjective night
(CT18.5), hamsters received an injection
of 300 nl of artificial CSF (aCSF)
containing either 1 mM CNQX (a competi-
tive, non-NMDA type EAA antagonist), 1 mM
MK-801 (a non-competitive, NMDA type
antagonist) or no drug (vehicle control)
directly into the suprachiasmatic hypo-
thalamus using a 33 gauge infusion can-
nula. Five minutes after injection, each
animal was exposed to light (30 lux) for
10 minutes. After light exposure, the
hamsters were returned to their cages and
maintained under DD for an additional 10
days. The effect of treatment on the
phase of the free-running activity rhythm
was determined as described previously
(Daan and Pittendrigh, 1976) using the
onset of wheel running activity as a
phase reference point. Injection sites
were verified histologically and only
data collected from animals with injec-
tion sites within 0.5 mm of the SCN were
included in the analysis.

In the absence of drug injection, light
treatment at this time results in a phase
advance of the free running activity
rhythm of approximately 81 ± 8 minutes.
Both drugs attenuated light-induced phase
advances by more than 85%. This result
suggests that EAA antagonists inhibit
light-induced phase shifts of the crica-
dian pacemaker by acting directly on the
SCN, possibly at the RHT synapse, and
support a role for EAA as the RHT neuro-
transmitter.

Light-induced Gene Expression

Light-induced resetting of the circa-
dian pacemaker represents a permanent
alteration in pacemaker function. The
c-fos protooncogene has been implicated
in the process of stimulus-transcription
coupling in the CNS (Curran and Morgan,
1987; Sagar et al., 1988) and appears to
mediate long-term adaptive changes in
neuronal function (Berridge, 1986; Goelet
et al., 1986). Recent work in our own
laboratory (Rea, 1989; 1992; Rea and
Michel, 1990) and elsewhere (Kornhauser,
et al., 1990; Rusak et al., 1990; Aronin
et al., 1990) has demonstrated that light
exposure during the subjective night
induces the expression of certain
immediate-early genes, including c-fos,
among a population of SCN cells.
Furthermore, light-induced c-fos expres-
sion only occurs when the light is
administered at a circadian time at which
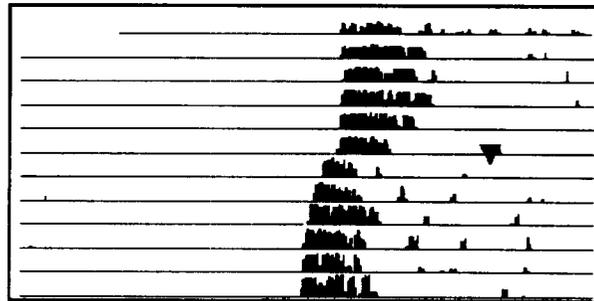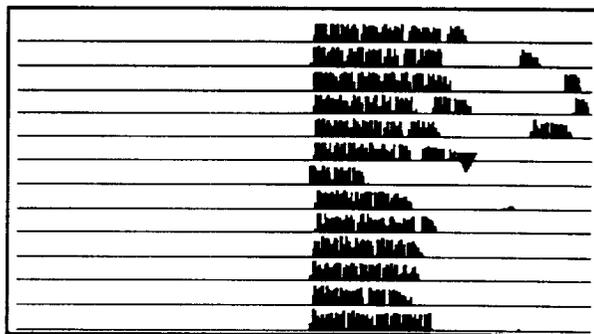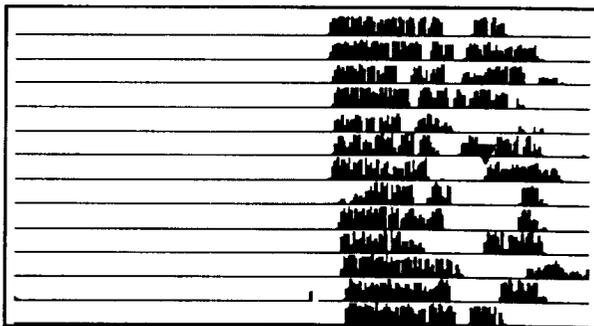
Actograms showing the effects of micro-
injections of vehicle (top), 1 mM CNQX
(middle), or 1 mM MK-801 (bottom) on the
light-induced phase advance of the free-
running activity rhythm. In all cases, a
brief light pulse (30 lux for 15 minutes)
was given at CT18.5 (inverted triangle).

478

a phase shift of the pacemaker results (Kornhauser et al., 1990; Rea, and Michel, 1990; Rea, 1992). These findings suggest that c-fos expression may represent a transcriptional event in the response cascade leading to light-induced phase alterations of the pacemaker. If this is the case, then ould expect that light-induced phase shifts and c-fos expression would share similar pharmacology.

This hypothesis was tested using our cannulated hamster model. Hamsters were housed in LD 14:10 for at least 7 days after surgery. At lights-out on the day before the experiment the animals were transferred to DD. Thiry hours after transfer (i.e., mid subjective night) the animals received microinjections of either aCSF or a solution of EAA antagonist in aCSF (300 nl at 1 mM) directly into the SCN. Five minutes later, each animal was exposed to 30 lux of white light for 10 minutes and returned to their cages under DD. Two hours after the onset of the light pulse, animals were deeply anesthetized, perfused transcardially with 0.4% paraformaldehyde, and the brains were processed for c-fos immunocytochemistry as described previously (Rea, 1989).

Light stimulation induces c-fos expression among a population of approximately 1055 $\pm$ 110 cells in the SCN. Both CNQX and MK-801 reduced the number of SCN cells that expressed c-fos in response to photic stimulation by about 45%. This result strenthens our proposal that c-fos expression represents an early transcriptional event mediating the effects of light on the SCN circadian pacemaker.

Significance of this Work to Aerospace Operations

The global mission of the U. S. Air Force demands that its personnel remain prepared at all times to participate in activities which are vital to the national security and likely to involve transmeridian air travel and irregular work schedules for sustained periods. In responding to this challenge, Air Force personnel are uniquely vulnerable to the performance limitations imposed upon them by the circadian timing system.

Knowledge gained from an investigation of the neurobiological basis of circadian rhythmicity will provide the database necessary for the development of photic and/or pharmacological strategies for alleviating the performance decrements associated with work schedules and practices which are incompatable with the circadian timing system. The elucidation of the mechanism of photic entrainment is an important step toward a detailed

understanding of circadian processes.

Pharmacological agents with specific and predictable effects on the circadian pacemaker could serve as useful tools for the control of circadian rhythmicity. For example, it may be possible to develop a specific antagonist against the RHT neurotransmitter. Such an agent could be used to selectively "blind" the circadian clock to the entraining influence of environmental light. This would alleviate the potential conflict between a shift worker's work-rest cycle and the environmental LD cycle. Similarly, pharmacological aids could be developed which would permit the rapid resetting of the circadian clock, accelerate rates of reentrainment of overt rhythms after rapid transmeridian travel, and maintain synchrony among shift workers in environments which provide unreliable or conflicting temporal photic cues, such as the proposed space station.

References

Albers, H. E.; Ferris, C.F., Leeman, S. E., Goldman, B. D. Avian pancreatic polypeptide phase shifts hamster circadian rhythms when microinjected into the suprachiasmatic region. Science 233: 833-835; 1984.

Albers, H.E.; Liou, SY.; Stopa, E.G.; Zoeller, T. Interaction of colocalized neuropeptides: Functional significance in the circadian timing system. J. Neuroscience, in press.

Aronin, N.; Sagar, S.M.; Sharp, F.R.; Schwartz, W.J. Light regulates expression of a Fos-related protein in rat suprachiasmatic nuclei. Proc. Nat. Acad. Sci. USA 87:5959-5962; 1990.

Berridge, M. Neurobiology: Second messenger dualism in neuromodulation and memory. Nature 323:294-295; 1986.

Cahill, G.M.; Menaker, M. Kynurenic acid blocks suprachiasmatic nucleus responses to optic nerve stimulation. Brain Res. 410:125-129; 1987.

Card, J. P.; Moore, R. Y. Ventral lateral geniculate nucleus efferents to the rat suprachiasmatic nucleus exhibit avian pancreatic polypeptide-like immunoreactivity. J. Comp. Physiol. 206:390-396; 1982.

Cassone, V. M.; Chesworth, M. J.; Armstrong, S. M. Dose-dependent entrainment of rat circadian rhythms by daily injection of melatonin. J. Biol. Rhythms 1:219-229; 1986.

Colwell, C. S.; Ralph, M. R.; Menaker, M. Do NMDA receptors mediate the effects of

light on circadian behavior? Brain Res. 523:117-120; 1990.

Curran, T.; Morgan, J.I. Memories of fos. Bioessays 7; 1987.

Daan S.; Pittendrigh, C.S. A functional analysis of circadian pacemakers in nocturnal rodents: II. The variability of phase response curves. J. Comp. Physiol. 106:253-266; 1976.

DeCoursey, P. J.; Buggy, J. Circadian rhythmicity after neural transplant to hamster third ventricle: specificity of suprachiasmatic nuclei. Brain Res. 500:263-275; 1989.

Earnest, D. J.; Sladek, C. D. Circadian rhythms of vasopressin release from individual rat suprachiasmatic explants in vitro. Brain Res. 382:129-133; 1986.

Earnest, D. J.; Turek, F. W. Role for acetylcholine in mediating effects of light on reproduction. Science 219:77-79; 1983.

Earnest, D. J.; Turek, F.W. Neurochemical basis for photic control of circadian rhythms and seasonal reproductive cycles: Role for acetylcholine. Proc. Nat. Acad. Sci. USA 82:4277-4288; 1985.

Ehret, C.F.; Potter, V. R.; Dobra, K. W. Chronotypic action of theophylline and pentobarbital as circadian zeitgebers in the rat. Science 188:1212-1215; 1975.

Gillette, M. U.; Reppert, S. M. The hypothalamic suprachiasmatic nuclei: Circadian patterns of vasopressin secretion and neuronal activity in vitro. Brain Res. Bull. 19:135-139; 1987.

Goelet, P.; Castelluci V. F.; Schacher S.; Kandel E. R. The long and short of long term memory -- a molecular framework. Nature 322:419-423; 1986.

Green, D. J.; Gillette, R. Circadian rhythm of firing rate recorded from single cells in the rat suprachiasmatic brain slice. Brain Res. 245:198-200; 1982.

Hebb, C.O.; Silver, A. Choline acetylase in the central nervous system of man and other mammals. J. Physiol. (London) 134:718-728; 1956.

Inouye, S. I. T.; Kawamura, H. Persistence of circadian rhythmicity in a mammalian hypothalamic "island" containing the suprachiasmatic nucleus. Proc. Natl. Acad. Sci. USA 76:5962-5966; 1979.

Ichikawa, T.; Hirata, Y. Organization of choline acetyltransferase-containing structures in the forebrain of the rat. J. Neurosci. 6:281-292; 986.

Johnson, R. F.; Morin, L. P.; Moore, R. Y. Retinohypothalamic projections in the hamster and rat demonstrated using cholera toxin. Brain Res. 462:301-312; 1988a.

Johnson, R. F.; Moore, R. Y.; Morin, L. P. Loss of entrainment and anatomical plasticity after lesions of the hamster retinohypothalamic tract. Brain Res. 460:297-313; 1988b.

Keefe, D.L.; Earnest, D.J., Nelson, D.; Takahashi, J. S.; Turek, F.W. A cholinergic antagonist, mecamylamine, blocks the phase shifting effects of light on the circadian rhythm of locomotor activity in the golden hamster. Brain Res. 403:308-312; 1987.

Kim, Y.I.; Dudek, F.E. Antagonism of fast excitatory postsynaptic potentials in suprachiasmatic nucleus neurons by excitatory amino acid antagonists. Neurosci. Abst. 15:1088; 1989.

Kornhauser, J.M.; Nelson, D.E.; Mayo, K.E.; Takahashi, J.S. Photic and circadian regulation of c-fos gene expression in the hamster suprachiasmatic nucleus. Neuron 5:127-134; 1990.

Lehman, M. N.; Silver, R.; Gladstone, W. R.; Kahn, R. M.; Gibson, M.; Bittman, E. L. Circadian rhythmicity restored by neural transplant. Immunocytochemical characterization with the host brain. J. Neurosci. 7:1626-1638; 1987.

Liou, S.Y.; Shibata, S.; Iwasaki, K.; Ueki, S. Optic nerve stimulation induced increase of release of 3H-glutamate and 3H-aspartate but not 3H-GABA from suprachiasmatic nucleus in slices of rat hypothalamus. Brain Res. Bull. 16:527-531; 1986.

Meijer, J. H.; Rietveld, W. J. Neurophysiology of the suprachiasmatic circadian pacemaker in rodents. Physiol. Rev. 69:671-707; 1989.

Moore, R.Y.; Lenn, N. J. A retinohypothalamic projection in the rat. J. Comp. Neurol. 146:1-14; 1971.

Newman, G. C.; Hospod, F. E. Rhythm of suprachiasmatic nucleus 2-deoxyglucose uptake in vitro. Brain Res. 381:345-350; 1986.

Pickard, G. E. The afferent connections of the suprachiasmatic nucleus of the golden hamster with emphasis on the retinohypothalamic projection. J. Comp. Neurol. 211:65-83; 1982.

480

Ralph, M. R.; Foster, R. G.; Davis, F. C.; Menacker, M. Transplanted suprachiasmatic nucleus determines circadian period. Science 247:975-979; 1990.

Ralph, M. R.; Menaker, M. Bicuculline blocks circadian phase delays but not advances. Brain Res. 325:362-365; 1985.

Ralph, M. R.; Menaker, M. Effects of diazepam on circadian phase advances and delays. Brain Res. 372:405-408; 1986.

Ralph, M.R.; Menaker, M. GABA regulation of circadian responses to light I. Involvement of GABA[A], benzodiazepine and GABA[B] receptors. J. Neurosci. 9:2858-2865; 1989.

Rea, M. Light increases Fos-related protein immunoreactivity in the rat suprachiasmatic nuclei. Brain Res. Bull. 23:577-581;1989.

Rea, M. Different populations of cells in the suprachiasmatic nuclei express c-fos in association with light-induced phase delays and advances of the free-running activity rhythm in hamsters. Molecular Brain Res. submitted.

Rea, M.; Michel, A.M. Phase-specific distribution of light-induced fos immunoreactivity in the SCN. Neurosci. Abst. 16:771; 1990.

Rusak, B.; Robertson, H.A.; Wisden, W.; Hunt, S.P. Light pulses that shift rhythms induced gene expression in the suprachiasmatic nucleus. Science 248:1237-1240; 1990.

Rusak, B.; Zucker, I. Neural regulation of circadian rhythms. Physiol. Rev. 59:449-526; 1979.

Sagar, S. M.; Sharp, F. R.; Curran, T. Expression of c-fos protein in brain: metabolic mapping at the cellular level. Science 240:1328-1331; 1988.

Sawaki, Y.; Nihonmatsu, I.; Kawamura, H. Transplantation of the neonatal suprachiasmatic nuclei into rats with complete bilateral suprachiasmatic lesions. Neurosci. Res. 1:67-72; 1984.

Swanson, L. W.; Cowan, W. M.; Jones, E. G. An autoradio-graphic study of the efferent connections of the ventral geniculate nucleus in the albino rat and the cat. J. Comp. Neurol. 156:143-164; 1974.

Shibata, S.; Liou, S.Y.; Ueki, S. Influence of excitatory amino acid receptor antagonists and of baclofen on synaptic transmission in the optic nerve to the suprachiasmatic nucleus in slices of rat hypothalamus. Neuropharmacology 28:403-409; 1986.

Takahashi, J. S.; DeCoursey, P. J.; Bauman, L.; Menaker M. Spectral sensitivity of a novel photoreceptive system mediating entrainment of mammalian circadian rhythms. Nature 308:186-188; 1983.

Turek, F. W.; Losee-Olsen, S. A benzodiazepine used in the treatment of insomnia phase-shifts the mammalian circadian pacemaker. Nature 321:167-168; 1986.

van den Pol, A. N.; Tsujimoto, K. L. Neurotransmitters of the hypothalamic suprachiasmatic nucleus: immunocytochemical analysis of 25 neuronal antigens. Neuroscience 15:1049-1086; 1985.

Zatz, M.; Brownstein, M.J. Injection of alpha-bungarotoxin near the suprachiasmatic nucleus blocks the effects of light on nocturnal pineal enzyme activity. Brain Res. 213:438-442; 1981.

Zatz, M.; Herkenham. M.A. Intraventricular carbachol mimics the phase shifting effects of light on the circadian rhythm of wheel-running activity. Brain Res. 212:234-238; 1981.

Acknowledgements

PHOTIC EFFECTS ON SUSTAINED PERFORMANCE

French[1],J., Whitmore[1],J., Hannon[2],P.J., Brainard[3],G., and Schiflett[1],S.

1 Armstrong Labs/CFTO Brooks AFB, TX USA 78235-5300
2 Health Sciences, Northern Arizona Univ. Flagstaff, AZ
3 Neurology, Jefferson Medical College, Philadelphia, PA

ABSTRACT

The advent of space exploration requires attention to the adaptability of human circadian rhythms in the unique environment of space. Circadian disruption, related to altered sleep work cycles and accelerated solar clues, can lead to fatigue that may impede mission success particularly as the duration of space flights increase. Research is described which evaluates manipulating environmental light intensity as a means to attenuate nocturnal fatigue.

A counter-balanced, within subjects design was used to compare 9 male subjects exposed to dim (100 lux) and bright (3000 lux) light conditions. Oral temperature values were greater for the bright light group over the dim light condition. Melatonin levels were suppressed by bright light treatment. Also, the frequency of eye blink rate was less for subjects during bright over dim light exposure. Light exposure was without effect on subjective fatigue. However, irrespective of light condition, significant effects on confusion, fatigue and vigor mood dimensions were found as a result of 30 hr sleep deprivation. The findings suggest that bright lights, may be used to help sustain nocturnal activity otherwise susceptible to fatigue. Such findings may have implications for the lighting arrangements on space flights during the subjective night for astronauts.

Key Words: Light, Temperature, Melatonin, Performance, Eyeblink, Mood

INTRODUCTION

Acclimatization to extraterrestrial environments represents a challenge to human productivity during future space missions. As extended flights become more frequent, a greater demand on the sustained vigilance of the crew increases the likelihood of performance problems associated with cumulative fatigue. Disrupted sleep is reported to be a common difficulty on shuttle flights particularly when dual shifts are required (Santy, et al., 1988). Fatigue problems related to alterations in circadian sleep work cycles and from unfamiliar light and dark solar cues have been known for some time and have been termed desynchronosis (Winget, et al., 1984). Although most astronauts quickly adjust to the demands of new work shifts in space, some never do and become chronically fatigued (Graeber, 1987). It may be that this fatigue results from an inability to resynchronize to the new circadian work rest cycles required in orbit. Light may serve as an adaptive counter measure for pre-shifting astronauts.

The adaptive characteristics of the circadian cycle to the unique environment of space is relatively unknown. Phase shifting of the sleep cycle as a result of travel across time zones on the earth can produce changes in the topography of the normal sleep EEG (Endo, et al., 1985) that may account for the inadequacy of the rest experienced during orbit. The fatigue produced by circadian disruption can be studied on earth albeit in the absence of microgravity. Effective treatment for fatigue related to circadian disruption may improve the potential for a successful mission. The present study evaluated the effects of ambient light as a counter measure to human fatigue degraded performance and may serve as a model of inducing circadian dysrhythmia.

Recent evidence supports a relationship between environmental light and improved nocturnal alertness (Campbell, et al., 1990, French, et al., 1990). The effectiveness of light exposure on performance enhancement is hypothesized to be related to the ability of light to attenuate the normal nocturnal surge of the pineal hormone melatonin. In support of this hypothesis, many studies suggest that melatonin acts as an endogenous sleep enhancing substance. For example, human subjects given relatively low doses (2 mg) of melatonin for three weeks experienced increased fatigue (Arendt, et al., 1984). Similarly, Lieberman, et al., (1985) using an acute oral dose of

240 mg of melatonin found reduced vigor, elevated fatigue, increased confusion and slowed reaction time. Additionally, plasma levels of melatonin are greatest during the sleep phase of the human circadian cycle. Orally administered melatonin has also been found to alleviate transcontinental disruption of circadian sleep wake cycles (Petrie, et al., 1989). Further, melatonin has a high affinity for receptor sites in the suprachiasmatic nucleus (SCN) of the hypothalamus where it is purported to trigger hormonal entrainment and regulate circadian and circannual rhythms (Reppert, et al., 1988; Brainard, et al., 1988).

Bright, light acutely suppresses plasma melatonin levels in animals (Benshoff, et al.,1987; Brainard, et al.,1982) including humans (Lewy, et al., 1980). The current study addressed the consequences of melatonin suppression via elevated ambient light intensity on temperature, melatonin levels, cognitive abilities, eye blink rate as measured by the electroogulogram and subjective mood.

METHODS

A counter-balanced, within subjects analysis of variance design was used to compare 9 male subjects exposed to dim (100 lux) and bright (3000 lux) conditions. Subjects were recruited from non-smoking civilian and military personnel who indicated a normal nocturnal sleep pattern. During both conditions, scores on cognitive performance tests developed for military human performance labs (Hegge, et al., 1985) were obtained every 2 hours throughout the 30 hour testing session. Beginning at 0600, subjects were stabiliized on the performance measures under dim light training conditions. Then at 1800, the light treatment (either dim or bright) began and continued until 0600 the next day. Finally, dim illumination was used until the completion of the experiment at 1200.

Immediately after each performance trial, oral temperature was measured and plasma samples were obtained for later melatonin assays (Brainard, et al., 1991). Monopolar electrodes attached to the bony orbit of the left eye and referenced to the pinna of the left ear were used to determine the blink rate for each subject during a 2 minute recording session, which also followed the performance trial. Blink rate per minute was then visually appraised from 1 minute of artifact free record in a blinded manner. Subjects completed profile of mood surveys (POMS) every 4 hours. They were then allowed 2 weeks before exposure to the second light condition. Subjects were prevented from drinking any caffeinated beverages and were fed the same foods (crackers, chips, sandwiches, fruit, pizza, milk, water, juices) at the same times during each light session.

Five subjects were evaluated at a time during each light session Each subject was assigned to a testing booth that contained a wide spectrum Vita-Light fluorescent lamp (Duro-test Corp., Fairfield, N.J., 07007 Part # 1157030) as the adjustable illumination source and a PC workstation. Each booth was separated from adjacent booths by sound attenuating, frame partitions that restricted the subject's view to their individual workstation. A comfortable chair allowed the subject to sit close to a work table that contained the workstation. The light source was mounted on a wooden frame over the workstation and suspended from an adjustable height to provide directed illuminance within either the dim or the bright treatment ranges. The subjects required to stay in the booth throughout the study with the exception of short (< 10 minute restroom breaks). Social interaction was kept to a minimum between subjects by the experimenter and by the demands of the testing schedule. Dependent measures on the cognitive tasks consisted of response time and accuracy variables. The order that the tests were presented did not vary throughout the study. The 10 performance tests used consisted of the choice reaction time (CRT), column addition and subtraction (CAS), the manikin test (MT), serial addition and subtraction (SAS) and Wilkinson reaction time (WRT). A subjective mood survey was also taken. As well, a tower puzzle (TP), following directions (FD), the numbers (N) and words (W) dual process task and route planning (RP) tests were used.

RESULTS

Oral temperature levels were significantly elevated in the bright light condition compared to the dim light condition at the 2130, 0130 and the 0330 sample points (p < .05) as shown in Fig. 1. Subjects in constant dim light had typically low levels of melatonin during daytime and higher levels at night. In contrast, this melatonin rhythm was suppressed by the bright light condition (Figure 2).
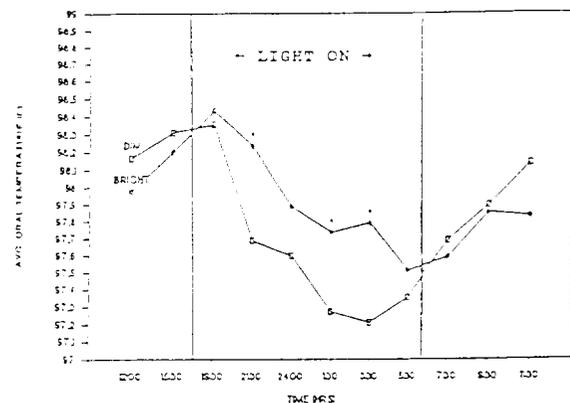
Figure 1. The effect of bright light on oral temperature. The onset of the bright light is indicated and differences from the dim light condition are marked with an *

483

NORMAL MALE VOLUNTEERS (N=7)
SUSTAINED PERFORMANCE
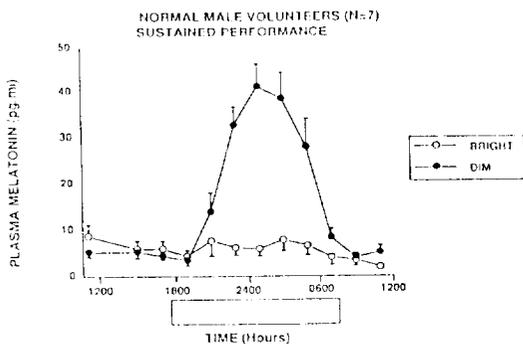
Figure 2. The suppression of normal nocturnal plasma melatonin levels was accomplished by bright light exposure. The open bar beneath the abscissa indicates the light exposure period.

Analyses of performance data indicate that bright light treatment improved response time while reducing the number of errors, particularly at the 2400 through the 0400 sample points. Table 1 shows the number of cognitive tests which were increased or decreased during bright light treatment. As shown in Table 1 the most effective time for bright light exposure to affect the cognitive tests occurred at midnight, 0200 and 0400 hrs. The effectiveness of the bright light exposure on cognitive ability did not extend beyond 0400 and may be detrimental when dim illumination is reinstated after 0600 as indicated in Table 1.

Table 1. The number of cognitive test results either increased or decreased by the application of bright lights at sequential times. A total of 10 performance tests were given at each time. The individual tests are identified (parentheses) at each time point. If response time or accuracy or if both were affected are indicated by a - or + or ±, respectively.

TIME	# OF TESTS INCREASED				# OF TESTS DECREASED	
Light on						
1800	1	(CRT+)			0	
2000	1	(RP+)			0	
2200	1	(SAS-)			0	
2400	4	(SAS-	N+	W+ FD±)	0	
0200	2	(SAS-	N+)		0	
0400	2	(SAS-	N+)		0	
0600	0				1	(FD±)
Light off						
0800	0				3	(FD± N+ W+)
1000	1	(SAS-)			1	(FD±)

Seven of the tests (CRT, RP, SAS, N, W, FD and MT) were sensitive to the effects of the illuminance conditions. Light exposure seemed to have beneficial effects on SAS, RP and N tasks as shown in Table 1, whereas the FD task seemed to be the most susceptible to disruption following extended exposure to bright light. Only response time performance variables were improved for the SAS test throughout the test session while the FD and W tasks demonstrated alterations in response time and accuracy variables at the times indicated. Only accuracy variables were susceptible on the N, CRT and the RP tasks. All of these results represent light x time awake interaction effects ($p < .05$). The bright light condition improved an accuracy variable (number of errors) on the MT as a main effect across all time points.

EYEBLINK FREQUENCY
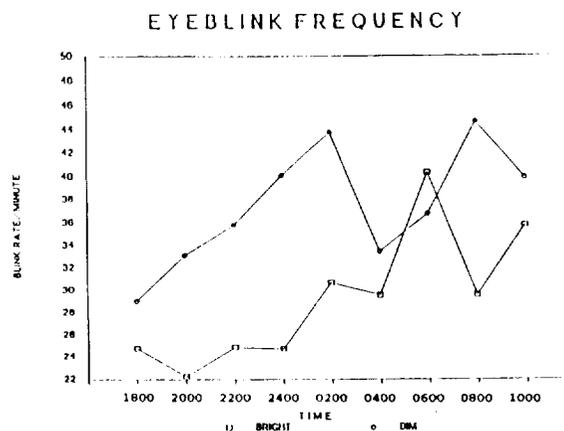
Figure 3. The increase in blink frequency per minute is shown during dim light (1800 - 0600 hrs) compared to the bright light exposure.

The results shown in Figure 3 demonstrates that light treatment was associated with significant differences in eyeblink rate as determined by the EOG. An overall main effect of light on eyeblink frequency was found ($p < .05$) but no interaction of light condition by time was found. However, the EOG differences did not parallel the time course of the melatonin or performance variables sensitive to bright light exposure.

There was no effect on subjective mood as a result of bright light exposure. However, as the duration of the sustained performance task increased 3 mood dimensions were affected as shown in Fig. 4. Subjective impressions of confusion, fatigue and vigor as shown in Fig 4a, 4b and 4c, respectively were affected during the later trials (time) when compared to the earlier trials, independant of light condition. Accordingly, the bright and dim light groups were averaged together in Fig. 4 for each mood state and graphed across hours awake. There was no effect on subjec-

484

tive anger, tension or depression as a result of extended hours awake, as shown in Fig. 4d, 4e and 4f, respectively.
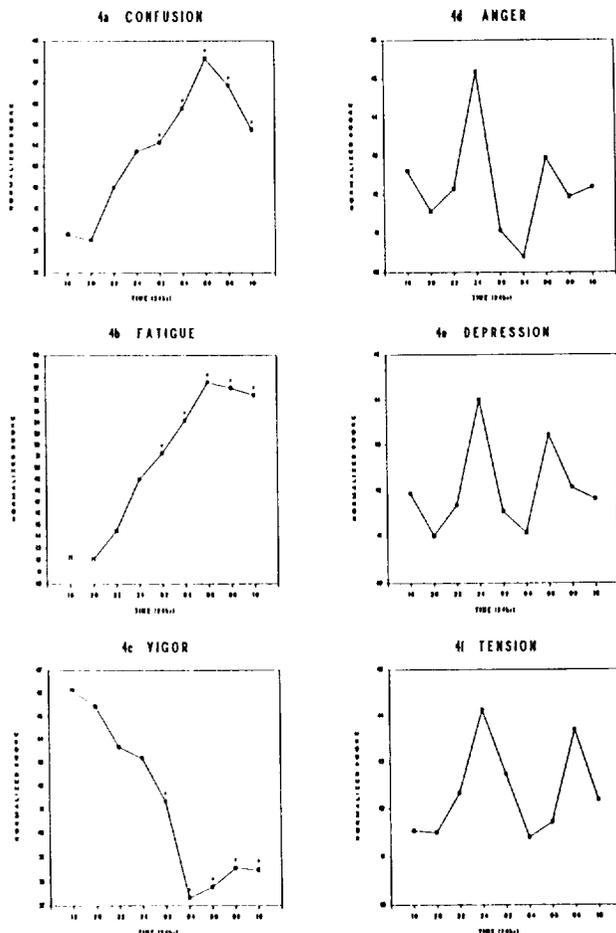
Figure 4. The effect of sustained performance on mood. No effects were found between light exposure conditions. Early trials were significantly different from later trials (p < .05) as indicated by an *. Bright and dim light groups were combined due to an absence of a light effect.

DISCUSSION

Exposure to bright light produced effects on oral temperature, melatonin levels and eye blink frequency. Although not sensitive to light condition, specific subjective mood states were responsive to the sustained performance test battery and to sleep deprivation. The times during which bright light exposure improved cognitive performance was similar to the time in which oral temperature was elevated and melatonin was suppressed. The levels of illumination used (3000 lux) were completely effective in suppressing melatonin to daytime levels. Although not excessive, 3000 lux seems to be more than adequate to control the normal nocturnal surge of melatonin.

Bright light exposure may improve performance otherwise susceptible to fatigue. However, it appears that there is no duration of the light effect beyond the exposure period. In fact, since performance begins to degrade somewhat after 10 hours of bright light exposure, the effectiveness of the lights in reducing fatigue degraded performance may have been exceeded. The absence of an interaction between light exposure and mood suggests that bright light did not improve mood state or make the the subjects feel less tired or more vigorous. Although effects on physiological state and cognitive performance were found, subjective mood was more sensitive to the duration of the sleep deprivation inherent in the 30 hour sustained performance test.

CONCLUSIONS

The protocol used may present the opportunity to evaluate adaptation problems, such as determining the optimal conditions for pre-shifting astronauts in earth bound labs. These and other problems associated with cumulative fatigue and shiftwork in the unique habitats and working conditions required in space could be studied with much greater facility in the absence of microgravity and the best solutions could then be applied to space operations. Also, the results may have implications for lighting conditions on board space flights, particularly for shift workers required to work during their subjective nights. Accuracy and response time might be improved by increased light intensity.

REFERENCES

Arendt, J., Borbely, A., Franey, C; & Wright, J. (1984) The effects of chronic, small doses of melatonin given in the late afternoon on fatigue in man: a preliminary study. Neuroscience Letters. 45:317-321.

Benshoff, H.M., G.C. Brainard, M.D. Rollag and G.R. Lynch,1987, suppression of pineal melatonin in Peromyscus leucopus by different monochromatic wavelengths of visible and near-ultraviolet light (W -A). Brain Research. 420:397-402.

Brainard, G., Richardson, B.A., Petterborg, L.J. and Reiter, R.J., The effect of different light intensities on pineal melatonin content, Brain Res., 233, 75-81, 1982.

Brainard, G., Lewy, A.J., Menaker, M., Fredrickson, R.H., Miller, L.S., Weleber, R.G., Cassone, V., and Hudson, D. Dose response relationship between light irradiance and the suppression of plasma melatonin in human volunteers, Brain Research, 454, 212-218, 1988

Brainard, G.C, Rollag, M.D., Hannon, P., French, J. and Storm, W.F., Effects of bright illumination on plasma melatonin in normal volunteers during sustained performance, Sleep Research, 20, 444, 1991.

Campbell, S.S. and Dawson, D. Enhancement of nightime alertness and performance with bright ambient light. Phys. and Beh., 48, 317-320, 1990.

Endo, S. and Sasaki, M. The possible mechanisms of the disturbed circadian sleep wake rhythm after time zone changes. Sangyo Ika Daigaku Zasshi, 7, Suppl. p 151-61, 1985.

French, J., Hannon P. and Brainard, G. Effects of bright illuminance on human performance and body temperature. Ann. Rev. of Chronopharm., 7, 45-49, 1990.

Graeber, R.C. Sleep in space. Proceedings of the 27th DRG seminar sleep and its implications for the military. p 59-69, 1987.

Hegge, F.W. Unified tri-service cognitive performance assessment battery (UTC-PAB). Design and specification of the battery. JWD3 MILPERF Report No. 85-1, U.S. Army Medical Research and Development Command, Ft. Detrick, Maryland, 1985

Lewy, A.J., Wehr, T.A., Goodwin, F.K., Newsome, D.A.,& Markey,S.P. (1980). Light suppresses melatonin secretion in humans. Science 210: 1267-1269.

Lieberman, H.R., Garfield, G., Waldhauser, F., Lynch, H.J., & Wurtman, W. (1985). Possible behavioral consequences of light-induced changes in melatonin availability. In The Medical and Biological Effects of Light. R. Wurtman, M. Baum, & J. Potts Jr., Eds. The New York Academy of Sciences. New York, 453:242-251.

Petrie, K., Conaglen, J.V., Thompson, L. and Chamberlain, K. (1989) Effect of melatonin on jet lag after long haul flights. British Medical Journal, 298, 705-707.

Reppert, S.M., Weaver, D.R., Rivkees, S.A. and Stopa, E.G. (1988) Putative melatonin receptors in a human biological clock,Science, 24, 78-79.

Santy, P.A., Kapanka, H., Davis, J.R. and Stewart, D.F. Analysis of sleep on shuttle missions. Aviat. Space Environ. Med., p 1094-7, 1988.

Winget, C.M., DeRoshia, C.W., Markley, C.L. and Holley, D.C. A review of human physiological and performance changes associated with desynchronosis of biological rhythms., Aviat. Space Environ. Med., 55 (12), p 1085-96, 1984

I

THE EFFECTS OF MULTIPLE AEROSPACE ENVIRONMENTAL STRESSORS ON HUMAN PERFORMANCE

S.E. Popper, D.W. Repperger, K. McCloskey
Armstrong Laboratory, Crew Systems Directorate
Wright Patterson Air Force Base, OH   45433-6573

L.D. Tripp
Systems Research Laboratories, Inc.
Dayton, OH

ABSTRACT

An extended Fitts' law paradigm reaction time (RT) task was used to evaluate the effects of acceleration stressors on human performance in the Dynamic Environment Simulator (DES) at Armstrong Laboratory, Wright-Patterson AFB, OH. The DES is a 19 foot radius man-rated centrifuge. This effort was combined with an evaluation of the standard CSU-13 P anti-G suit versus three configurations of a "retrograde inflation anti-G suit" (RIAGS) manufactured by the David Clark Company. Seven subjects participated in four (4) "blend" runs and four (4) data runs on the centrifuge. The 4 blend and data runs corresponded to the number of anti-G suits evaluated (1 standard and 3 RIAGS). A blend run consisted of the initial combining of the RT task with G_z acceleration for each suit configuration. A data run was identical to a blend run, but it was assumed subjects were now familiar with the experimental set-up. Each run consisted of the following acceleration profiles: 1) a 4 G_z warm-up for 15s, 2) a 1 minute rest at baseline (1.4 G_z), and 3) a modified simulated aerial combat maneuver (SACM) consisting of +4 G_z to +7 G_z alternating plateaus, each 15 seconds in length. The SACM was performed until peripheral light loss (PLL); physiological discomfort occurred (usually due to anti-g suit configuration), or fatigue. Results indicated that RT and error rates increased 17% and 14% respectively from baseline to the end of the SACM and that the most common error was pressing too few buttons.

INTRODUCTION

Reaction Time Task

The modeling of the human information processing system using reaction time (RT) techniques dates back over 100 years ago to the work of the Dutch physician Donders (4,8,13,14). Donders proposed that RT is a "composite" score that includes stages of perception/discrimination, a choice process, and a reaction from the subject. These three stages have usually been defined as occurring serially (12,14). The use of RTs as an index of human information processing is based on the concept which assumes "...the time from stimulus to response will be sensitive to the speed of the [central neurological] processing responsible for [response] selection..."(13).

RTs obtained from choosing between alternative stimuli came to be known as choice RT (14,15). The relationship between the choice RT and the number of stimulus alternatives were mathematically described as a Log_2 function by both Hick (6) and Hyman (7), known formally as the Hick-Hyman law.

> Choice RT = $a + b[log_2 (N)]$ (1)
> where N is the number of stimulus-response alternatives and *a* and *b* are empirical constants.

The Hick-Hyman law states that there is a linear relationship between the response time of the subject and the log_2 of stimulus alternatives. This highlights one of the major concepts contained within this law: it is assumed that the time required to make a decision about a response is linearly related to the amount of information needed to make that decision (4).

As we have made use of the term *information* earlier while describing what humans do (*human information processing*), some kind of definition seems warranted. Here, information is strictly defined as the amount of uncertainty that is reduced by the fact that a signal was presented.

The amount of information conveyed is a direct function of the amount of uncertainty prior to the presentation of the signal, as well as by the amount by which uncertainty is reduced. In general, the amount of information (H) is given by:

$$H = Log_2(1/P_i) \quad (2)$$

where P_i is the probability that a given event (i) will occur.

H is measured in *bits* where one *bit* is defined as the amount of information necessary to reduce the original uncertainty by half or one alternative of choice. Relating this to the Hick-Hyman law, every time the number of stimulus-response alternatives is doubled, the amount of information to be processed is increased by 1 bit (and presumably, choice RTs also increase by a *constant* amount).

In designing the choice RT task, two variables are important: (1) the nature of the relationship between the stimuli and the associated responses and (2) practice or experience with the task. The term, stimulus-response (SR) compatibility, is a measure of how natural the connection is between the stimulus and the response. The more natural the relationship between a stimulus and response, the less time required to process 1 bit of information (reflected in a smaller value of the slope of the RT function - b) and hence, an increased capacity of the human. The effect of practice may develop a high degree of compatibility between a stimulus-response pair normally considered incompatible.

It was our intent to develop a performance task that could easily discern changes in cognitive ability as the subjects were affected by the stressors. Such a performance task must be extremely sensitive to elicit changes due to the combinations of the various stressors acting on the subject.

One such task developed at the Armstrong Laboratory involves an extension of the classical Fitts' law paradigm in a multi-dimensional sense, which can be considered as a subset of the Hick-Hyman law (2). This type of task investigates the tradeoffs of speed to accuracy as humans perform simple and complex reaction time tasks. The Fitts' law paradigm is ideal for this research in the sense that it includes both a metric to evaluate task difficulty as well as a measure of capacity (or baud rate) in the accomplishment

of a task (in a temporal sense) as well as increase the amount of errors that occur.

Another advantage of using this extended Fitts' law paradigm is from the information contained in the errors. In the task developed in this study, four types of errors occur and they illustrate when (and under what circumstances) the task completion process breaks down. Analysis of these errors indicate "how" the capacity is compromised as the subjects are exposed to multiple stress situations.

The motivation for extending the Fitts' law paradigm in this paper is derived from the work of Agarwal, et al. (2). In this study, it was shown that by using multiple stimuli and responses, the task could be made more and more difficult to perform until finally the subject would break down and make a substantially larger number of errors. The manner in which the task was made more difficult was accomplished by presenting the stimuli at a faster and faster rate, thus producing a form of difficulty in a temporal sense. Task difficulty could also be increased by having larger numbers of stimulus response pairs in the task scenario.

The use of linear RT models to describe and evaluate human information processing capacities is not universally condoned; for example, generalizing RT results to complex human activities such as playing basketball or flying an airplane is at best incomplete (11). Nevertheless, RT methods have been used extensively in the past to quantify the effects of environmental stressors on human performance capabilities (1). As such, an RT method based on an extension of Fitts' law is used in the present study.

Attention

Another concept that must be dealt with, as it plays a major role in human performance, is attention. There are many definitions of attention, but most agree that it is sometimes serial, sometimes parallel, concentrated, limited, and focused. Attention is felt to have limitations in the capacity to handle information from the environment. This leads to the concept of interference where two tasks are performed simultaneously and the degree to which they interfere with each other are measured. If two tasks can be performed as well simultaneously as individually, then at least one task may not require attention and can be called *automatic*, or the tasks may be referred to as being independent in their access to certain types of processing resources. If there is some decre-

ment in the performance of a task when performed with another, then both tasks are considered *attention demanding*, and not independent in resource "drain." There are two types of interference; structural and capacity. Structural interference occurs when two demands are placed on physical/ neurological structures (i.e., requiring the hand to be at two places at the same time). If no structural interference exists, then a capacity interference is inferred. This inference is based on the assumption that there is a limitation to some central capacity resource (attention).

There are multiple theories explaining attention; undifferentiated, fixed capacity (single channel), flexible allocation, multiple resource, and functional view (as the result of a choice all other processes are prevented from occurring or only with great difficulty, (18)). Several mechanisms of parallel sensory processing have been described (11). The *Stroop phenomenon* occurs when the same stimulus in two different conditions is relevant, but in one of the conditions a secondary stimulus is processed at the same time causing an increase in the RT. A classic example is where subjects are to respond to the color (red, blue, green, yellow) of different geometric forms (triangles, circles, squares), versus responding to the colors of words which correspond to the colors (i.e., responding to the word 'blue' when printed in yellow versus responding to the printed color red when it appears as the word 'green').

As other examples of attention phenomena, the *dichotic listening paradigm* describes how man can ignore one of two messages presented through headphones. However, there are certain messages that cannot be ignored i.e., when your name is spoken. The *psychological refractory period (PRP)* states that the reaction time (RT) to the second of two closely spaced stimuli is considerably longer than RTs to the first stimuli.

A final area to address is the relationship between attention, stress, motivation, and arousal. Arousal or activation are usually considered neutral terms that describe the energy level of the individual. The term neutral is used because arousal represents the amount of effort being applied to whatever action is being accomplished. It can range from deep sleep to the highly energized state characteristic of an individual fighting for survival or competing in an important sporting event (13). Stress and motivation have a directional component where stress is considered negative while motivation implies movement towards a

goal (13). A classic relationship exists between arousal and performance as discovered by Yerkes and Dodson (20) commonly referred to as the *inverted-U hypothesis*. There is an optimum arousal level to obtain peak performance. Any more or less will cause a decrease in performance.

RT and Acceleration

What do these theories have to do with the present study? It is hypothesized that acceleration (the presence of a greater than 1 G stressor) will have a detrimental effect on a serial RT processing paradigm, specifically an extended Fitts' law processing task. Under acceleration, error rates should increase. RT should also increase. However, the RTs to each of the response conditions used below, one (1), two (2), or three (3) button choices, may or may not retain their differences. In other words, whereas under normal conditions the RTs increase as the number of button choices increase (a positive slope), this relationship may not hold under acceleration (a 'flat' slope) due to attentional resources being diverted to the task of maintaining physiological integrity under high-G (9).

METHODS

Task Equipment and RT

Figure 1 illustrates the RT device used by the subjects in this experiment. The stimuli presented were combinations of *one, two, or three out of four possible lights in this diagram. The subjects kept two fingers (the index and middle finger) of each hand on the four buttons. To complete the task, the buttons corresponding to the illuminated lights had to be pressed. RT was calculated as the time between the onset of the stimuli (lights) and the corresponding button presses. The stimuli were presented at different *interstimulus times* (t_i), where the time between the presentation of each stimulus was varied. The three values for this variable were 800 msec, 400 msec, and 200 msec. SR compatibility was considered high because of the spatial relationship between the "on-screen" stimulus lights and response buttons. Practice with the task at normal 1 G_z was accomplished until each subject was able to maintain greater than 80% accuracy.

FIGURE 1
Reaction Time Task

Lights
→
Stimuli

□ □ □ □

Buttons
→
Responses

○ ○ ○ ○
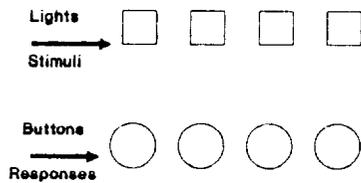
Errors

There were four classes of errors a subject could make:

(1) The subject could wait too long (any response more than two seconds after the stimulus onset was dubbed "sleep time").

(2) The subject had to press the buttons *simultaneously*. This meant that if a subject pressed any two or three keys more than 50 msec apart, an error was recorded. This helped the subjects to approximate a "simultaneous" response when more than one button was to be pressed.

(3) The subject could press the wrong number of buttons (either too few or too many), or the incorrect buttons.

(4) If a subject responded within 100 msec of stimulus onset, an "anticipation time" error was recorded. This was used to reject any responses smaller than the human choice RT limitation of approximately 160 msec.

Subjects

Subjects were four (4) males and three (3) females, aged 23 to 40 years, obtained from the Sustained Acceleration Panel (qualified subject pool). All subjects had undergone extensive medical screening before acceptance on the panel.

Experimental Variables

To fully describe the experiment, the forms of the environmental stressors (acceleration and anti-G suit configuration) need to be elaborated. The acceleration stressor and the anti-G suit configurations were independent variables, while an attempt was made to control for individual variability through randomization.

Acceleration: The modified simulated aerial combat maneuver (SACM) was selected as the acceleration stressor and represents a typical combat scenario.

Figure 2 displays the alternating 7.0 G_z to 4.5 G_z SACM profile. The term G_z refers to that component of the acceleration stressor acting from head to foot. The end result is decreased blood profusion (pressure) at head and eye level resulting in visual degradation and ultimately loss of consciousness unless steps are taken to maintain sufficient blood pressure at higher G_z levels. This is accomplished through the appropriate use of the anti-G straining maneuver (AGSM) in conjunction with an anti-G suit.

FIGURE 2
ACCELERATION STRESS CONDITION

Gz Stress

0:00 0:30 1:00 1:30 2:00 2:30 3:00 3:30 4:00 4:30 5:00 5:30 6:00 6:30
Time

— Series 1
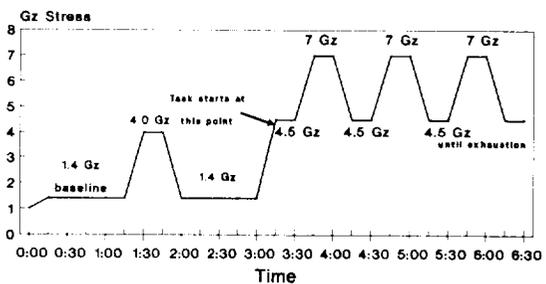
Modified SACM

Anti-G Suits: The experiment was also designed to evaluate several different configurations of anti-G suits and how they impacted subject performance. The standard anti-G suit in use today (the CSU-13P) is composed of 5 bladders that are inflated caudalward with pressure increasing linearly as G increases, compressing the abdomen, both thighs (quadriceps), and both calves. The retrograde inflation anti-G suit (or RIAGS) has the same bladder configuration but inflates cephaladward and is considered a *full-coverage* suit (has the appearance of a pair of pants, not *cutaway* as the standard). Perhaps the most uncomfortable aspect of the suit is abdominal pressure, which is an important factor providing protection when combined with leg pressure (19). The degree of discomfort is a function of the fit of the suit, placement of the abdominal bladder against the subject's diaphragm, and the individual's personal opinion .pa concerning increased abdominal pressure. The pressures maintained in these suits were approximately 8.5 psi at 7.5 G_z. These pressure levels are very uncomfortable at 1 G_z, but are tolerable at the higher G levels depending upon the individual. In addition to the full-coverage RIAGS, there were added two different

490

|

types of arm counterpressure, namely, occlusion cuffs and pressure sleeves. This arm counterpressure was assumed to reduce the amount of blood pooling into the arms during acceleration. However, each of the two arm configurations were based on different counterpressure techniques, specifically an arterial occlusion technique (cuff) and a more widespread counterpressure technique along the length of the arm (sleeve). In summary, four different types of anti-G suit configurations were used:

1) standard CSU-13 P anti-G suit
2) RIAGS alone
3) RIAGS with sleeves, and
4) RIAGS with cuffs.

The mechanisms used to explain the effectiveness of anti-G suits for human G-protection are: (1) anti-G suits increase peripheral resistance, thus improving eye-level blood pressure under G; (2) anti-G suits help prevent rapid extravasation of plasma from the blood vessels into tissue during G stress by offering immediate counterpressure; (3) anti-G suits may play a role in increasing venous return, particularly with simultaneous inflation of both leg and abdominal bladders; and (4) anti-G suits support and raise the diaphragm, thus mechanically supporting the heart and decreasing the heart-to-eye distance (3,19).

Obviously, the bottom line for protection is effective counterpressure during G (presumably, the more the better (19)). However, human factors issues must be taken into consideration (16,17). Increased abdominal pressure via the inflated bladders may cause discomfort even under high-G (19). There are wide individual differences concerning discomfort; some subjects are not bothered at all while others devote more energy trying to breathe during a high-G run than while performing the more strenuous AGSM.

Individual Variability; Withstanding High-G Until Exhaustion: This brings us to the issue of individual variability. The level of experience on the centrifuge is a factor in how the subject devotes attention to the AGSM while performing another task (9). The subject's level of G-tolerance also dictates how well he or she will be able to maintain the AGSM while concentrating on another task.

G-tolerance is a function of physical fitness, time elapsed from last G-exposure, and miscellaneous other factors contributing to the general stress level each subject experiences (5). There is a wide discrepancy in subject response to anti-G suits, as well as different experiences with peripheral light loss (PLL) (10). Much of the research in acceleration has used PLL as an objective measure of individual stress levels. When PLL reaches the point of a 60 degree cone around the central visual axis, the subject normally terminates the run. However, not all subjects have symmetrical PLL. Thus, when we say that a subject continued the SACM to exhaustion and include this as the "end-point" within our design, we implicitly assume wide individual differences in the definition of "exhaustion" (i.e, 60% PLL, abdominal and other bodily pain, fatigue, etc.).

Dependent Measures: Subjects were intrumented with arterial oxygen saturation (SaO_2) plethysmography (mounted on the earlobe), a transcranial Doppler (TCD) sensor mounted at the temple, and electrocardiographic (ECG) chest leads. Data collected were time at G until exhaustion and termination of the session, heart rate obtained from the ECG, the time course and level of SaO_2, the time course and level of blood velocity obtained from the TCD, error rate and type, RT, and subjects' ratings of suit comfort. However, only the RT and error rate data are reported here.

Experimental Design

During a complete session, subjects began with 180 "warm-up" practice trials on the RT task at 1 G_z (termed "pre acceleration"). This normally took less than five minutes to accomplish. Immediately following these practice trials, subjects were then accelerated to a baseline of 1.4 G_z for 1 minute, followed by a 4 G_z run for 15 seconds (which hopefully provided some physiological pre-adaptation to G). Subjects remained at baseline for 60 seconds after which time the SACM profile began (Figure 2). At the first 4.5 G_z peak, the task was presented and continued for the duration of acceleration until exhaustion, when the subject terminated the exposure (termed "peak acceleration"). In addition, the task continued for approximately 50 more trials after exposure so as to provide data during the recovery phase (termed "post acceleration").

A complete session was accomplished 8 times; 4 "blend" sessions and 4 data sessions were accomplished (one blend session and data session each for the four anti-G suit configurations outlined

above). The four blend sessions were completely randomized within subjects, as were the four data sessions. A blend session was the initial coupling of task performance with high-G. Data sessions were identical to the blend sessions, except subjects were now familiar with performing the task under high-G. All results reported below were obtained from the data sessions.

RESULTS

Initial evaluation of the data indicated that interstimulus times (200, 400, or 800 msec) had no significant effect on either RT or error rate/type, regardless of the number of stimuli (1, 2, or 3 lights/buttons). Thus, RT and error rate/type data were collapsed across interstimulus times in the results presented below.

RT

Our first analysis concerned the overall effect of the pre, peak, and post acceleration conditions on RT. All three conditions were significantly different from one another, $F(2,12) = 17.21$, $p < 0.0003$. RTs occurred in a descending order: peak was larger than post, which in turn was larger than pre (peak=504.4; post=464.4; pre=425.8).

Table 1 shows the statistical results of the effects of suit configuration and the number of lights/buttons on RT for each of the pre, peak, and post acceleration conditions. As can be seen, there were no interactions between suit configuration and lights/buttons for either the pre, peak, or post acceleration conditions. However, there was an effect of suit configuration on RT during post acceleration. RTs were longer for the RIAGS with cuffs configuration than for RIAGS with sleeves.

For the pre acceleration condition, there was a significant main effect for lights/buttons, where RT increased as number of lights/buttons increased (which of course was expected according to the Fitts' law paradigm, see Figure 3a). What is interesting is that this lights/buttons effect was not significant during peak acceleration (or more correctly, at exhaustion, see Figure 3b). The effect returned at post acceleration (Figure 3c).

To further examine this effect, we performed individual F-tests for each of the suit configurations for the pre, peak, and post acceleration conditions. Table 2 shows the statistical results. For pre acceleration, all four suit conditions show significant effects for lights/buttons, in the same pattern (3 is larger than 2, which in turn is larger than 1). For the peak acceleration condition, there were no significant differences for lights/buttons. At post acceleration, the effect returns, but with a difference. For the RIAGS alone and the RIAGS with cuffs conditions, RTs for 3 lights/buttons are significantly larger than for 1, but not from 2. The standard and RIAGS with sleeves show the same pattern as for pre acceleration (3 > 2 > 1).

Error Rate/Type

The types of error generated at peak (exhaustion) for the entire experimental design are shown in Table 3. As can be seen, sleep time errors occurred the least, while pressing too few buttons was the most common type of error. There were no errors where subjects pressed multiple keys more than 50 msec apart and are not shown in Table 3.

TABLE 1. The Effects of Suit Configuration and Number of Lights/Buttons on RT

ACCELERATION CONDITION:	PRE	PEAK	POST
TEST:*			
Suit*Button	NS	NS	NS
Suit	NS	NS	RIAGS/c > RIAGS/s F(3,36) = 5.31, p< .0085 490.8 > 420.4
Button	3 > 2 > 1 F(2,12) = 157.87, p < .0001 472.1 > 430.7 > 374.8	NS	3 > 2 > 1 F(2,12) = 90.13, p < .0001 508.1 > 463.9 > 420.4

* Suit*Button: suit configuration by lights/buttons interaction; Suit: suit condition main effect;
Button: lights/buttons main effect.

TABLE 2. The Effects of Acceleration Condition on Lights/Buttons RT by Suit Configuration

ACCELERATION CONDITION:	PRE	PEAK	POST
*TEST:			
STD	F(2,12) = 91.24, p < .0001 3 > 2 > 1 427.89 > 430.28 > 374.86	NS	F(2,12) = 80.10, p < .0001 3 > 2 > 1 504.95 > 453.07 > 413.15
RIAGS	F(2,12) = 308.51, p < .0001 3 > 2 > 1 468.86 > 428.47 > 375.18	NS	F(2,12) = 12.03, p < .0014 3 > 1 507.25 > 429.79
RIAGS/s	F(2,12) = 83.01, p < .0001 3 > 2 > 1 471.34 > 431.41 > 373.75	NS	F(2,12) = 32.52, p < .0001 3 > 2 > 1 477.87 > 443.83 > 398.39
RIAGS/c	F(2,12) = 92.08, p < .0001 3 > 2 > 1 475.47 > 432.54 >375.46	NS	F(2,12) = 13.10, p < .0001 3 > 1 542.31 > 440.44

* STD: standard anti-G suit; RIAGS: retrograde inflation anti-G suit; RIAGS/s: RIAGS with sleeves;
RIAGS/c: RIAGS with cuffs.

FIGURE 3

a PRE

b PEAK

c POST

| STD | RIAGS | RIAGS/c | RIAGS/s |

TABLE 3 - Type of Errors at Peak (Exhaustion)

Type*	Total
Sleep Time	3
Anticipation	12
Too many buttons	16
Wrong buttons	18
Too few buttons	65

* see text for explanation of error types.
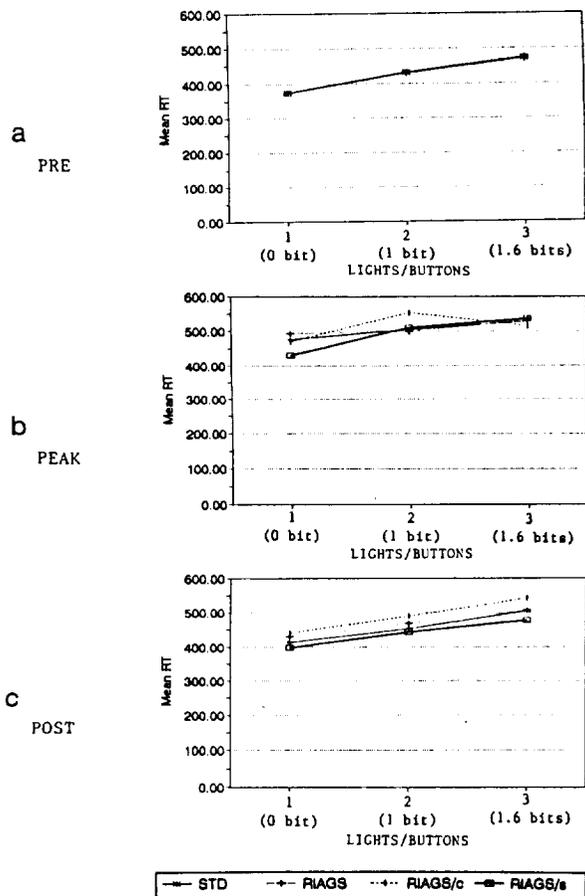
Total error rates at peak acceleration (from sleep time to inaccurate button presses) for each of the four suit configurations are shown in Table 4. The pre acceleration condition is used as a "baseline" here, and each of the error rates for the suit conditions were compared to this baseline. Error rates for the standard anti-G suit and for the RIAGS with cuffs conditions were significantly larger at peak acceleration than at pre acceleration. RIAGS alone and RIAGS with sleeves did not differ in error rate from pre acceleration.

Comfort and Time to Termination

Subjects' rankings of suit comfort, and their total time under high-G before exhaustion and termination of the exposure, are shown in Table 5. As can be seen, these variables match each other in terms of superior rankings; in short, the RIAGS with sleeves was ranked highest, followed by RIAGS alone, the standard suit, and RIAGS with cuffs.

TABLE 4 - Pre Acceleration Error Rates Compared to Peak for Each Suit Configuration.

	Pre Acceleration	STD*	RIAGS	RIAGS/c	RIAGS/s
ratio of errors to total	2/78	7/78	6/71	25/78	5/83
percent	3%	9%**	8%	32%**	6%

* STD: standard anti-G suit; RIAGS: retrograde inflation anti-G suit; RIAGS/c: RIAGS with cuffs; RIAGS/s: RIAGS with sleeves.
** $p < 0.05$, one-tailed t-test comparison of pre acceleration to peak for each suit configuration.

TABLE 5. Rankings of Suit by Comfort and Time Until Termination of Acceleration Exposure (Exhaustion).

RANKINGS

Suit:*	Most Comfortable	Most Time Under Acceleration
RIAGS/s	1	1
RIAGS	2	2
STD	3	3
RIAGS/c	4	4

* STD: standard anti-G suit;
RIAGS/c: RIAGS with cuffs;
RIAGS: retrograde inflation anti-g suit;
RIAGS/s: RIAGS with sleeves.

DISCUSSION

The original hypothesis that RTs would be longer under acceleration than at pre or post acceleration was supported here, which strongly suggests that high-G interferes with the human's central information processing capacity. In addition, the linear relationship between number of lights/buttons and RT (as lights/buttons increase, so does RT in a linear fashion) was not supported under high-G at subjects' exhaustion point. A possible reason for this finding could be that subjects were greatly preoccupied with their physiological and bodily integrity at the point of exhaustion and were devoting few attentional resources to the completion of the task; at exhaustion, subjects are at their physiological and psychological limit. They need to divert more of their attention from the task (stimuli) to maintaining head level blood pressure through the use of anti-G

straining maneuvers (AGSM) to prevent blackout and loss of consciousness. Other factors more difficult to quantify, and which may also serve to explain this effect, are the psychological consequences resulting from the situation subjects find themselves in (anxiety, fear, pain, ego, etc.). In future studies, some combination of these factors may fully explain the lack of attention given to the task under high-G at the point of exhaustion.

The type of suit configuration did not have an effect on RT, even though the comfort and time to termination rankings showed a definite pattern (the RIAGS with sleeves was superior to RIAGS alone or to the standard suit, while the RIAGS with cuffs seemed to be inferior). The error rates did show a corresponding pattern, however. The RIAGS suit with cuffs had the largest error rate of all suit conditions, while it was also ranked the most inferior in terms of comfort and time to termination. Most likely, this was due to the intense discomfort of the occlusion cuffs resulting in termination of the SACM secondary to numbness/pain rather than fatigue. It could be said that an additional stressor was added to the design matrix due to the nature of the cuffs. The cuffs occluded blood going to and coming from the lower arms. Over long periods of time, subjects reported that their arms would "go numb" and feeling would cease, or become an overriding "tingling pain" sensation.

Interestingly, the ranking of g-suits by error rate matches exactly the ranking of the suits by comfort as well as the total time to termination. The most comfortable suit had the least errors just prior to termination, as well as the greatest time under acceleration (RIAGS with sleeves).

In conclusion, acceleration stress indeed had an impact on the RT model. Because of the nature of the stressor, attention was diverted during peak Gz causing a loss in the ability to discriminate between stimuli responses. Implications to the Air Force in support of its mission are as follows: 1) how to quantify for each individual pilot the net effect of multiple stressors (physical fatigue, mental fatigue, length of sorties, type of profiles within a sortie, number of sorties per day) and predict the point at which peak Gz "exhaustion" occurs; and 2) what is the correlation of the fatigue status of pilots versus the sortie workload, both during high-g maneuvers and "normal" flight. These questions require further studies.

494

REFERENCES

1. Advisory Group for Aerospace Research and Development (AGARD, 1989), "Human performance and assessment methods" (Aerospace Medical Panel Working Group 12, AGARD/NATO-AG-308), AGARD/North Atlantic Treaty Organization: Neuilly Sur Seine, France.

2. Agarwal GC, Agrawal P, Corcus D, .lm5 Gottlieb GL, Repperger DW, "Cognitive performance measures VIA response equivocation", Proceedings of the 23 Annual Conference on Manual Control, 1987.

3. DeHart RL, ed. FUNDAMENTALS OF AEROSPACE MEDICINE, Philadelphia: Lea & Febiger, 1985.

4. Dupin, COURS DE GEOMETRIE ET DE MECANIQUE APPLIQUEE, cited from L. Walther, LA TECHNOPSYCHOLOGIE DU TRAVAIL INDUSTRIEL, Paris, 1926, pg 13.

5. Grissett JD, "Physical fitness to enhance aircrew g tolerance," USAF-SAM-SR-88-1, Naval Air Station, Pensacola, FL.: United States Air Force Scholl of Aerospace Medicine and the Naval Aerospace Medical Research Laboratory, 1987 Joint Service G-tolerance Conference. (NTIS No. ADA-204689).

6. Hick WE, "On the rate gain of information", QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 4, pg 11-26, 1952.

7. Hyman R, "Stimulus information as a determinate of reaction time", JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 45, pg 188-196, 1953.

8. Pieters JPM, "Sternbergs' additive factor method and underlying psychological processes: Some theoretical considerations," PSYCHOLOGICAL BULLETIN, 93, 411-426.

9. Ponomarenko VA, Oboznov AA, Arkhangel'skiy DYu, "Mental control of physiological state under effect of sustained longitudinal accelerations", KOSMICHESKAIA BIOLOGIIA I AVIAKOSMICHESKAIA MEDITSINA, 21(2):24-7, Mar-Apr, 1987. FTD-ID(RS)T-0112-90.

10. Popper SE, "Unequal narrowing of the visual field in a $+G_z$ environment", AVIAT. SPACE ENVIRON. MED., In Press.

11. Rumelhart DE, McClellan JL, PARELLEL DISTRIBUTED PROCESSING: EXPLORATIONS IN THE MICROSTRUCTURE OF COGNITION, Volume 2, Foundations, Cambridge, MA: The MIT Press, 1986.

12. Sanders AF, "Stage analysis of reaction processes, In G.E Stelmach and J. Requin (Eds.). TUTORIALS IN MOTOR BEHAVIOR, pg 331-354, Amsterdam: North-Holland, 1980.

13. Schmidt RA, MOTOR CONTROL AND LEARNING: A BEHAVIOR EMPHASIS, 2nd Edition. Human Kinetics Publishers, Inc., Champaign, IL. 1988, pg 75-139.

14. Sternberg S, "The discovery of processing stages: Extensions of Donder's method, In W.G. Koster (Ed.), ATTENTION AND PERFORMANCE II, Amsterdam: North-Holland, 1969.

15. Turvey MT, "Preliminaries to a theory of action with reference to vision", PERCEIVING, ACTING, AND KNOWING, R. Shaw and J. Bransford (Eds), 1977.

16. Viteles, MS, Editorial Foreward, "Human Engineering for an Effective Air Navigation and Traffic Control System," Paul Fitts (Ed), National Research Council, Air Navigation Development Board, March 1951.

17. Wiener EL, Nagel, DC, (Eds), HUMAN FACTORS IN AVIATION, Academic Press, Inc., San Diego, CA., 1988.

18. Wickens, CD, ENGINEERING PSYCHOLOGY AND HUMAN PERFORMANCE, Charles Merrill Publishing Co., Columbus, OH, 1984.

19. Wood EH, Lambert EH, Code CF, Baldes EJ, "Factors involved in the protection afforded by pneumatic anti-blackout suits," National Research council, Division of Medical Sciences, Committee on Aviation Medicine, Report No. 351, 24 Aug 44.

20. Yerkes RM, Dodson JD, "The relationship of strength of stimulus to rapidity of habit-formation," JOURNAL OF COMPARATIVE NEUROLOGY AND PSYCHOLOGY, 18, 459-482. 1908.

MICROGRAVITY EFFECTS ON STANDARDIZED COGNITIVE PERFORMANCE MEASURES

Samuel G. Schiflett, Ph.D.
Sustained Operations Branch
Armstrong Laboratory
Brooks AFB Texas 78235-5000

ABSTRACT

The purpose of this experiment, selected to fly on International Microgravity Laboratory (IML-2) spacelab mission, is to determine the effects of microgravity upon cognitive skills which are critical to successful performance of many tasks on board the space shuttle. Six tests from the Unified Tri-service Cognitive Performance Assessment Battery (UTC-PAB) will be administered to the Mission Specialists to fulfill the goals of this experiment. These tests are based upon current theoretical models of human performance, an the hypothesized effects of microgravity. The principal objective is the identification of the effects of microgravity upon specific information processing skills affecting performance from those of fatigue and shifts in work/rest cycles. Multiple measures of both short- and long-term fatigue will be obtained and used as a major independent variable for the analysis of these performance data. Scientific supporting studies entitled "Training schedules to acquire and maintain performance stability" will determine optimum practice and performance testing schedules for the astronauts. The same tests will be used post-flight to collect data on the recovery of any cognitive performance impairment compared to pre-flight, baseline levels.

BACKGROUND

Problem

Astronauts are subject to a variety of stresses during space flight. These stresses include microgravity, physical isolation, confinement, lack of privacy, fatigue, and changing work/rest cycles (Christensen & Talbot, 1986). Any one or a combination of these stressors could degrade the cognitive skills required to perform tasks essential to the success of the mission. Of these potential stressors, the effects of fatigue and of the changing work/rest cycle are known to cause deteriorations in astronaut productivity (Stepanova, 1975).

Purpose

The purpose of the inflight research is to determine the effects of microgravity on cognitive skills required by many tasks on board the space shuttle. This paper describes; (1) the performance tests and the rationale for selection, (2) the pre/post baseline and inflight experiment, (3) the supporting ground-based studies that will determine the reliability and stability of the measures, and (4) the controls necessary to differentiate the effects of fatigue and changing work/rest cycles from microgravity.

Previous Research

To date, no systematic attempt has been made to determine the effects of space flight on cognitive skills. Despite the lack of systematic observation, anecdotes and documented instances of degraded performance do exist (Covault, 1988). Additionally, two experiments have examined performance in space. The first (Ross, Schwartz, & Emmerson, 1987) demonstrated a deterioration in mass discrimination during flight, which persisted for approximately 3 days after touchdown. The authors did not pinpoint the source of the deterioration although some aspects of judgment and psychomotor coordination were implicated. The second (Ratino, Repperger, Goodyear, Potor, & Rodriguez, 1988), examined the simple reaction time, choice reaction time, and time perceptions before, during, and after a mission. The simple and choice reaction time tasks showed no effects of space flight. The time perception task showed an increasing deterioration in estimates of short durations (2 to 16 s) throughout the mission. Again, the locus of the deterioration was not identified.

It is anticipated that our experiment will benefit from having the results from the IML-1 space flight experiment, Mental Workload and Performance Evaluation (MWPE) scheduled to fly approximately two years before IML-2. The MWPE experiment will evaluate the most effective human computer interface while astronauts perform a "Fittsberg" task (Hartzell, Gopher, Hart, Lee, & Dunbar 1983). The task will simultaneously measure the accuracy and reaction time of retrieving items from short term memory and skilled motor coordination response times while using three different types of input/output controllers; joystick, keyboard, and trackball. The

C-2

differential effects of microgravity on the output stage of human information processing will be particularly beneficial to our study since it will use the same hardware system available to us at the Performance Assessment Workstation (PAWS) on board the space shuttle. Even though additional measures of subjective workload and mood state will be taken from the Mission Specialists during the MWPE, IML-1 flight, the primary objective of the "Fittsberg" task will yield psychomotor response times to evaluate human engineering issues (Newman & Bussolari, 1990).

The "Fittsberg" task combines the more common Serial Memory Search Task (Sternberg, 1969) and the historical discrete motor movement Fitts Task (Fitts & Peterson, 1964). Even though the task combination was designed for another purpose, it is important to realize that if generalizations or predictions are to be made of astronaut performance from one space flight mission to another, with varying task demands and durations; then consideration should be given to standardizing the testing procedures and using more common transportable system software. These factors will become increasingly important when large performance databases must be shared to accomplish future international space flights. Therefore, due to the scarcity of reported cognitive performance data, it is mandatory that additional well controlled, cognitive studies using "standardized" procedural testing, be conducted before long term space flights are flown. This paper outlines the origin of such a performance test development program by the military that is relevant to NASA operations.

Human Performance Testing in Military Medical Research

For many years military medical researchers have recognized that the development of microcomputer-based tests of human performance would have broad applicability. The resulting test systems employed a variety of tests, batteries, and hardware, all purporting to measure such functions as memory, information processing abilities, logical reasoning, tracking, etc. (Bittner, Carter, Kennedy, Harbeson, & Krause 1984; Shingledecker, 1984; Thorne, Gensen, Sing, & Hegge 1985). While these test systems were suitable for answering isolated questions; standardization of procedures, software language, and hardware was lacking, making comparison of results from different laboratories difficult. Several years ago, in response to the challenge of evaluating the side effects of chemical defense pretreatment and antidote drugs on human performance, the U.S. Army Medical Research and Development Command formed the Joint Working group for Drug Dependent Degradation of Military Performance (JWGD3 MILPERF). This tri-service working group was broadly tasked with the mission of developing standardized methodologies for the assessment of human performance.

The UTC-PAB

One of the principal products of the JWGD3 effort has been the Unified Tri-Service Cognitive Performance Assessment Battery (UTC-PAB), a set of 25 standardized human performance tests which run on a microcomputer (Englund et al., 1987; Hegge et al., 1985; Perez et al., 1987; Reeves et al., 1989). The tests comprising the UTC-PAB were selected after an exhaustive search of the human performance literature and in-depth interviews with the original test authors. A team of experienced human performance investigators evaluated the candidate tests for validity, reliability, and sensitivity to provide investigators and practitioners with the means to sample a broad range of human performance functions. The prototype "beta testing" phase of the UTC-PAB has been completed. It includes a detailed set of hardware-independent specifications, as well as an authoring system to construct additional prototype tests for concept formation. The original 25 tests run on IBM-compatible microcomputers using the system clock, standard system keyboard, and a serial port joystick for tracking tests. This NASA project will implement several of the tests, discussed later in this paper, on a NASA supplied GRID 1530 computer. Additional implementations of the UTC-PAB tests in a variety of forms are currently under way, including a version of the tests for clinical neuropsychologic screening that will run on several laptop computers. All of these activities are being coordinated by the Office of Military Performance Assessment Technology (OMPAT) located at the Walter Reed Army Institute of Research (WRAIR) in Washington, DC, which assumed the functions of the JWGD3 in August 1989. The founder of the triservice JWGD3 and the current Director of OMPAT is Dr. Fred Hegge.

The UTC-PAB has become recognized worldwide as a standard for military performance testing. The AGARD NATO (Working Group 12) has recently published the test specifications of a subset of the UTC-PAB tests for research with environmental stressors (AGARD, 1989). The STRES battery is an example of a product of international standardization coordinated by OMPAT. The tests were selected by NATO, specified by the Air Force, and programmed by the Navy. Through this cooperative effort the OMPAT is continuing the process of establishing a mechanism for collecting the data from numerous international studies using a subset of the UTC-PAB. The existing data base at OMPAT has grown into a central Performance Information Management System (PIMS) located at WRAIR in Washington, D.C. This networked data base will enable the establishment of international performance norms for the UTC-PAB tests, as well as document the effects of a wide variety of environmental stressors and drugs on human performance. Free access to the PIMS data base and performance measurement tools is granted by the Director, OMPAT at a toll-free, bulletin board number 1-800-542-7844.

GENERAL APPROACH

In order to capitalize on the standardization already achieved in performance testing by the Department of Defense and NATO, the Life Sciences Division at NASA Headquarters competitively selected a team of researchers responding to a NASA Headquarters research announcement. A portion of this announcement was initiated by Dr. Janis Stoklosa, Chief, Human Factors, Behavior, and Performance, to study the effects of microgravity on astronaut performance on IML-2. The team is headed by the author of this paper, the Principal Investigator, Dr. Samuel Schiflett. The government co-investigator is Dr. Jonathan French, from the Sustained Operations Branch of the Armstrong Laboratory Brooks AFB, Texas (formerly the Crew Performance Function of the USAF School of Aerospace Medicine). Other participating co-investigators are Dr. Douglas Eddy, NTI, Inc. and Dr. Diane Damos, University of Southern California.

Six tests from the UTC-PAB were selected by the team, based on current theoretical models of human performance and on the hypothesized effects of microgravity. These tests will be further studied in ground-based environments to determine required parameters for use in space. Final parameters will be incorporated into software that will run on the specified Space Shuttle hardware, and within the allowed time constraints.

The actual experiment to be performed in space will involve testing each Mission Specialist for 20 minutes a day using behavioral and subjective tests selected to evaluative cognitive functioning while in orbit. Particular emphasis will be given to the question of fatigue. Multiple measures of both short- and long-term fatigue will be obtained and used as a major independent variable in analyzing the performance data.

RATIONALE FOR SELECTING TESTS

Several factors must be considered in selecting tests for use on board the space shuttle. One of the most important of these is the restrictive time available during flight for performance assessment. Another constraining factor is the prohibitive cost of developing flight worthy hardware not previously furnished by NASA. A critical factor to consider is the specific information processing skills necessary to the success of the mission. The final issue that is the most relevant to this experiment, is the information provided by a specific test should aid in identifying the cognitive processes or information processing stages affected by microgravity. These four factors were taken into account in selecting the six tests included in the performance assessment battery.

Many different cognitive skills are critical to the successful completion of tasks on board the shuttle. Upon examination of past Mission Specialist's tasks and insight gained at a recent NASA workshop held at JSC on Human Factors requirements, and from conferences such as this one; a list of functional task areas were identified to be potentially affected by microgravity. Of these functions, spatial information processing, tracking, and time-sharing play a disproportionately important role in the success of a variety of Spacelab and Spacestation tasks. For this reason, a tracking test, a spatial information-processing test, a time-sharing test, and a directed attention switching test have been included in the battery. Two other tests, the Sternberg Memory Search test and the Continuous Memory test, have been included because each of these examines specific cognitive processes and stages of information processing. Thus, the locus of microgravity effects, if any, on cognitive functioning can be identified relatively clearly. A description and rationale for each test is given below. Detailed descriptions and procedures may be found in Perez et al. (1987) for all the tests and combinations except the Matrix test and the Manikin-Mathematical Processing combination. Simple and choice reaction time tasks were not included in the proposed battery even though they have reliable and short training times because Ratino et al. (1988) failed to identify any performance deterioration during space flight.

NASA PERFORMANCE ASSESSMENT BATTERY (PAB)

After careful review of the statistical properties of each test and examining theoretical rationale of the UTC-PAB, the following performance tests were selected for use in this experiment.

Tracking. One of the primary potential effects of microgravity is a disruption of visual-motor coordination due to disturbances in the sensory input and motor output channels. To probe for this effect, a tracking test will be used. One of the candidate tracking algorithms is the Crossover Model developed by McRuer and Jex, (1967) and validated by DOD, NASA, FAA in numerous studies. It requires the subject to maintain a target in the center of a horizontal line. A fixed difficulty (lambda level) is used to displace the target, and the subject must manipulate a control device to null this input disturbance. Even though, it has limitations on its approach (frequency domain), two characteristics recommend it for inclusion in the proposed battery. First, performance reaches differential stability in 150 brief trials (Damos, et al., 1984). Second, it is known to be affected by a variety of exotic environments, including alcohol (Klein & Jex, 1975; Dott & McKelvy, 1977), hypoxia (Nesthus, Schiflett, Bomar & Holden, 1988), fatigue (Gevins, Cutillo, Fowler-White, Illes, & Bressier 1988).

However, in the present experiment, several tracking algorithms must be evaluated prior to specifying the tracking test's final configuration. Some of these issues related to pilot workload and dynamics have been addressed in a conference sponsored by the Air Force (Frazier & Crombie, 1982). Alternate models will be considered such as the Optimal Control Model (Levison, Barron, & Kleinman, 1969) that is a time-domain

498

approach to establishing operator describing functions. Alternate tracking algorithms will be evaluated prior to the ground-based support experiments described in a later section of this paper. However, the tracking task will be presented at a difficulty level empirically derived for each Mission Specialist slightly below the maximum resource capacity, i.e. subcritical mode, if the Unstable Tracking task is used. A compensatory tracking task may be more diagnostic to subtle changes in psycho-motor processing (Wickens, 1986).

Matrix Rotation Task. The functional de-afferentation of the otolith organs that occurs in space, along with associated effects on the visual and cerebellar systems, raise serious questions about the individual's spatial processes. Thus, it is desirable to probe for subtle effects on the person's ability to perceive, remember, and process spatial information. The Matrix Rotation task developed by Phillips, (1974) and Damos & Lyall, (1984) will be used to evaluate the effects of microgravity on spatial processing. The Matrix Rotation Task should not be confused with the task implemented by Thorne, et al. 1985 or the commonly used Match-to-Sample task reported by Thomas & Schrot, (1988) at the Naval Aviation Medical Institute. This test uses 100 basic patterns. Each pattern is a 5 by 5 matrix with five illuminated cells that have been selected at random. At the beginning of the trial, the subject sees a pattern. After the subject studies the pattern, he presses a response key. The pattern is immediately erased and a new one presented. The subject must decide as quickly as possible if the new pattern is identical to the preceding pattern. The subject then presses one key for "same" or another key for "different." As soon as the response is made, a third pattern appears. The subject must now compare the new pattern to the immediately preceding pattern, etc. For "same" responses, the two patterns are never presented in exactly the same orientation; the second pattern is always rotated either 90 degrees to the left or 90 degrees to the right relative to the preceding pattern. Both correct reaction time and percentage correct are used as dependent measures.

Although no data were located in the open literature which examined the effects of exotic environments on this test, it does have two advantages. First, performance on this test reaches differential stability in approximately 15 minutes. Thus, little practice is required. Second, concurrent verbal suppression tasks have been shown not to affect performance on the Matrix task, indicating that this task indeed measures some aspect of spatial information processing. Studies examining the effect of fatigue and changes in the sleep rest cycle using this task are under development by the military.

Sternberg Memory Search. The general Sternberg paradigm requires subjects to respond as rapidly and accurately as possible to visually presented letters. At the beginning of the test a set of letters drawn randomly from the alphabet are presented to the subject for memorization. The set of letters (positive set) stay on the screen for 10 seconds, then the screen is cleared and a series of single test letters are presented. If the presented letter matches one of the letters in the previously memorized positive set, the subject responds "same" (key press). If a different letter appears (negative set), then the subject responds "different" (key press) indicating a non-matching letter was presented.

Some questions have been raised about the reliability of the intercept and slope scores derived from the mean reaction time of multiple size memory sets (Carter, Krause, & Harbeson, 1986). Reliability is a concern if the fixed set procedure for presenting the stimuli is used for a large number of trials (Wickens, Vidulich, Sandry, & Schiflett 1981). In this procedure, an item that has been designated as a target in one trial can never be used as a distractor (non-target) in subsequent trials. Thus every time a target stimulus is presented, the subject should respond. This presentation procedure results in a gradually decreasing slope that, with a high number of trials and sufficient practice, will be statistically indistinguishable from zero. The Sternberg task included in this version of the UTC-PAB uses a set size of four letters that are changed after each block of trials. Thus, a letter can be a target in one session and a distractor on another. Carter, et al. (1986) found that a variable four letter memory set represents the more cognitive aspects of information processing, is both reliable and stable, and is highly correlated with the other letter set sizes (1,2, or 3). This will be particularly helpful in evaluating the microgravity effects on the cognitive comparative processing stage. Using a single set size of four letters will save time and provide data on the locus of subtle deleterious effects of the microgravity environment on performance.

Continuous Recognition. One critical aspect of higher cognitive function is the ability to maintain attention and to carry out repetitive cognitive processes over some period of time. In many ways, such activities encompass those which were traditionally referred to as "vigilance." However, they add the dimension of active processing of information, rather that simple monitoring. One task that appears to capture the performance elements above is the Continuous Recognition Test (Hunter, 1975; Shingledecker, 1984). In this the subject sees two numbers, one above the other. The task is to remember the bottom number. When the next two numbers appear, the task is to determine if the new top number is the same as the previous bottom number. However, before responding, one must note the new bottom number because as soon as a response is made, the numbers are replaced by a new pair. Thus, the subject must not only exercise very short-term memory, but more importantly, must inhibit the response until the new bottom number is committed to memory. The appropriate strategy is to develop a set pattern of observing, memorizing, observing, comparing, and responding. This sequence is different enough from that

required by most routine tasks that it requires constant attention allocation. Even brief lapses result in errors. The task can be made even more difficult by requiring the subject to remember and respond to numbers further removed from the immediately preceding one (e.g., two- or even three-back), thus imposing a much higher load on immediate memory.

This task has not been used as often as others in the UTC-PAB battery. However, it has a respectable data base indicating that it is a sensitive test of both short-term memory and attention allocation (Perez et al., 1987). It provides a large number of data points in a short period of time, and has good test-retest reliability.

Prior to the inflight experiment, it will be necessary to develop additional data on the parameters that will optimize the value of the data collected in space. This will be done in the context of the preliminary studies described in other sections of this paper. We will determine, for instance, the optimum amount of practice required for stability, as well as the value of using the one-, two-, or three-back condition. The continuous recognition test is known to be sensitive to g-stress (Ross & Chambers, 1967) and to alcohol (Carpenter & Ross, 1965).

Dual Tasking--Tracking and Sternberg Memory Search. One of the most critical and potentially sensitive higher cognitive functions that might be affected by microgravity is the ability of the subject to allocate attentional resources among several tasks. To investigate this, the present study will use the time-sharing paradigm that has been well studied in cognitive psychology (Damos & Wickens, 1980; O'Donnell & Eggemeier, 1986; Damos, (1991). The specific form of this paradigm will be the dual task included in the UTC-PAB. This consists of the Sternberg task and the Tracking task being presented simultaneously. In this implementation of the Dual Task, the tracking task is presented in the middle of the screen and the letters of the Sternberg task appear in a fixed location directly above the center null point. The target of the compensatory tracking task moves laterally. One memory set will be used, consisting of 4 letters. Due to the nature of the dual task, the "fixed set" procedure must be used, in which the same memory set letters are presented with several probe letters for each daily session. For a recent study discussing the implementation of the dual task when investigating the effects of antihistamines on military weapon system controllers see Nesthus, et al. 1991).

Performance Switching Task--Manikin and Mathematical Processing. Time-sharing, as explained above in the Dual task, is different from another required attentional process that could be affected by microgravity. Astronauts must make rapid shifts in the attentional focus, as well as in the skills required to respond to a change in task demands. This externally-directed behavior defies automaticity in any true sense, since it must be flexible

enough to respond to unusual demands. Thus, a test is needed to probe the subject's ability to shift attention and resource allocation in response to rapidly changing and unpredictable external demands. Such a procedure has been created that uses two tasks currently in the UTC-PAB.

In this procedure, the subject has two distinct and discrete tasks to perform. One is a spatially-based task, and the other is a mathematically-based task. Each of these appear, trial by trial, simultaneously on the screen. However, an arrow appears at the same time directing the subject which task is "active" (i.e., must be responded to). The subject must make an exclusive response to the active task, where reaction time and percent-correct data are obtained only for that task. The switching from task to task for each trial is random (within constraints). Therefore, the subject must remember to watch the arrow on each trial, allocate the appropriate resources to respond to that trial, and then make the appropriate response. This paradigm provides a test of the attention switching skills described above.

The two tests selected to exercise this paradigm are the Manikin test and the Mathematical Processing test. The Manikin test has a long history of use (Benson & Gedye, 1963; Reader, Benel, & Rahe, 1981;) and is presented in a wide variety of formats by military psychologists (Miller, Takamoto, Bartel, & Brown, 1985). As implemented in this microgravity experiment, a manikin "stick figure" is presented facing either forward or backward. In addition, the figure can be either upright or upside-down. The figure is also standing on a box and inside the box is either a rectangle or a circle. In the figure's two hands are a rectangle and a circle. The subject's task is to note which symbol is inside the box, and then to determine which of the manikin's hands is holding the designated symbol. The subject then presses the left or right of two keys corresponding to the manikin's left or right hand.

Exposing the Mission Specialists to a series of manikins presented in a variety of orientations in microgravity where the astronaut has the freedom to position his own body during work and sleep in non-traditional gravity orientations; will be one of the more theoretical interesting tests of spatial processing using a human form. The effects of microgravity on whole body orientation will give insight into whether this test is a perceptual measure of spatial transformation of mental images or involves pure spatial abilities requiring readaptation (Carter & Wolstad, 1985).

The Mathematical Processing test is based on similar tasks described by Perez et al. (1987). It presents two single-digit numbers that must be added or subtracted. If the answer is greater than 5, one response is given. If the answer is less than 5, another response is required. This task has been reported by Shingledecker

500

(1984) to be a relatively pure index of mathematical functioning.

METHODS/TECHNIQUES

The software for the NASA Performance Assessment Battery (PAB) is being programmed and delivered under the oversight of Dr Samuel Moise, NTI, Inc. in collaboration with Ms Kathryn Winter, Naval Aerospace Medical Research Laboratory. The software team is following a software verification plan that has been mutually agreed upon by NASA and the Air Force. The software will be in executable form, on 3.5" floppy disk, compatible with the GRID 1530 microcomputer and the attached controller devices; joystick or trackball. The programs will reside on the hard disk prior to flight and be backed-up with system disks and individual data storage disks for rapid loading during flight, if required. Provision for entering the final values on the tracking task for each astronaut will be made available if late access (Launch -3 days) is granted.

All system integrity checks and the initial transfer of the programs to the GRID hard disk are automatically performed by software provided by the Air Force. Start-up of the test system requires only a single command typed by the Mission Specialist at the GRID's keyboard. The six performance tests, a subjective fatigue scale, and a mood questionnaire will be sequenced by a control program that will automatically present all tests once started.

Data from all tests will be stored on disk in a format compatible with the UTC-PAB normative data base. It is estimated that the data will occupy approximately 35K of disk space per Mission Specialist per day. For the entire 13 day mission, this would be a total of 1.61 megabytes for all the raw data to be collected by these tests. This is well within the storage capacity of the on-board computer system. As a back-up, the data will also be stored in summary statistical format on each individual 3.5 inch data disk. The data files will be formatted using the UTC-PAB and NATO standards for data processing by commonly used statistical software packages.

RESEARCH PLAN

Phase I Ground-Based Scientific Support Studies

The primary purpose of the first series of ground-based scientific support studies is to obtain three types of data related to reliability, training, and normative database. The first type provides basic information about the trial-to-trial reliability of the tests and the amount of practice necessary to reach differential stability. The second type will provide data on the most efficient training scheme. These data are necessary to ensure that all Mission Specialists are completely trained before launch. The third type will provide a normative data base on all of the tests. This data base will be collected on approximately 100 subjects with the appropriate range of age and male/female representation. This first series of studies will only be conducted after the test battery has been tested on the final NASA Grid computer configuration. Dr Diane Damos, Co-investigator, University of California will be responsible for all aspects of the study while using the coordinated technical expertise of the USAF Sustained Operations Branch staff and contractors from NTI, Inc. These data will supplement the large existing UTC-PAB data base that exists for most of these tests.

A control study, sponsored by the Air Force, will be conducted in the sustained operations laboratory to demonstrate the sensitivity of the tests to fatigue. The performance battery will be identical to the one used in the test reliability and stability study. Approximately eight to twelve subjects will be selected as closely as possible from the same general population as the mission specialists. Each subject will be tested under both rested and fatigued conditions separated by at least one week. Order of testing will be counterbalanced. Several physiological and biochemical covariate measures will be taken (Brainard, Hannon, French, & Storm 1990; Brainard, Rollag, Hannon, French, & Storm 1990).

Fatigue has been manipulated in several ways within the sustained operations laboratory (French, Hannon, & Brainard, 1989; Morris, 1984). To demonstrate the sensitivity of the proposed battery to fatigue, sleep deprivation will be manipulated. The SAM Fatigue scale, the Sleep Quality Inventory, and the Mood Scale which have demonstrated sensitivity to fatigue in a wide variety of research and long-term field studies will be used to supplement the performance test results (Storm, Dowd, & Boll, 1990).

Phase II Flight Experiment

The Phase II flight experiment will test several hypotheses relating to the interactive effects of microgravity, work/rest shifts, and resulting fatigue on cognitive performance. All the hypotheses will be tested against the null hypothesis that there is no difference among the groups, nor across days. Each test- and scale-dependent measure will be evaluated separately.

The IML-2 flight will consist of two work/rest cycles shifted by 12 hours. A typical shift schedule is shown in Figure 1. It is expected that two of the Mission Specialists assigned to Shift A will gradually be time adjusted forward 12 hours prior to flight over a period of several days. This Shifted Group will go to sleep after a few hours in orbit when the Non-Shifted Group (two other Mission Specialists) assigned to Shift B will complete their duty day.
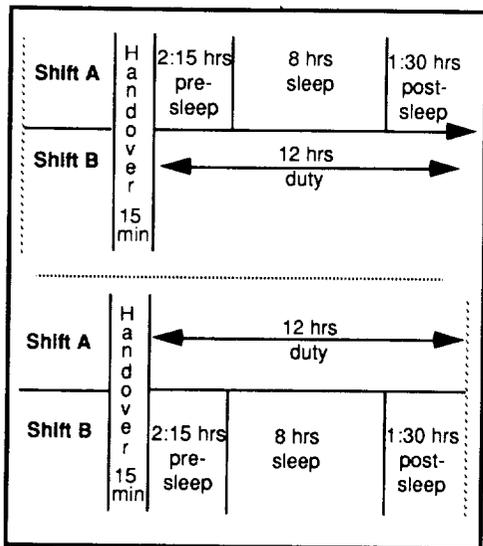
Figure I. Typical Two-Shift On-Orbit Work/Rest Cycle

The two major independent group variables include microgravity and the effects of ground phase shifting. The groups will be as follows:

GROUP	PRE-BASELINE	FLIGHT	POST-BASELINE
	-LI8 to -L3	MD1 to MD13	+R1 to +R16
Non-phase shifted	*0800-2000	0800-2000	0800-2000
Phase shifted	partial	2000-0800	0800-2000

*assume 0800 Launch L=Pre-Launch MD=Mission R=Recovery Day

TABLE 1 - PERIODS OF DATA ACQUISITION FOR EACH GROUP

General Hypotheses

1. Performance will deteriorate across days as the result of time in microgravity relative to the preflight baseline for both groups. Cumulative fatigue will increase concurrently.

2. Performance will deteriorate more across the first 3-5 days for the phase shifted groups relative to the non shifted groups. Fatigue will increase, then decrease in a similar manner.

Specific cognitive skills affected by microgravity will be indicated by the specific tests effected. For example, longer reaction times to the four item Sternberg Memory Search Test would indicate effects on central memory scanning processes.

The scores on tests and subjective scales during the preflight testing will provide a baseline for each subject. Difference scores for each subject will be derived by subtracting pre-flight baseline scores from flight scores.

Any residual effects of in orbit space flight will be evaluated using post-flight scores obtained in the recovery phase of the experiment. All hypotheses will be tested on the portional baseline difference scores using an analysis of variance statistical model. The 13 days of the inflight test will be a within group independent variable.

Phase III Post Flight - Mission Simulation Control Study.

Although the research team members believe that substantial insight into understanding the effects of microgravity and fatigue on performance can be gained using the proposed flight paradigm, serious consideration should be given to sponsoring a Phase III, Mission Simulation Control Study. This proposed study which is presently not funded by the Life Sciences Division of NASA is still at the planning stages. The data from a group of ground-based subjects performing tasks as similar as possible to those in space in the same sequence would provide the best comparison group for isolating the effects of microgravity from those of fatigue and circadian rhythms on performance. The drastic impact on performance data and fatigue ratings of a shuttle emergency or other unanticipated event can not be duplicated without such a post flight mission simulation control group. Mission related interruptions during sleep could be sufficient to cause Mission Specialists to experience fatigue that could deteriorate performance on a long mission such as IML-2.

If funded, the Mission Simulation Control Study would manipulate all the variables of the proposed flight study, except microgravity; including work/rest cycle shifts and simulated daily routine tasks involving exercise and housekeeping chores, e.g. meal preparation. The purpose of this study will be to demonstrate the sensitivity of the measures under conditions similar to the Phase II flight test. These data would serve as a comparison data set for the Phase II flight. The location of this study has not been determined but preliminary arrangements for mutual scientific exchange visits with Dr. Alex Gundel from DLR/DARA, Cologne, Germany has begun. Dr. Gundel has been identified by NASA as a participant to acquire data from the a IML-2 Life Sciences Experiment conducted by Dr. Timothy Monk, University of Pittsburgh entitled "Human Sleep, Circadian Rhythms, and Performance in Space". Dr. Gundel will use the sleep and circadian rhythm data from IML-2 to reconstruct, through computer simulation, a model of the IML-2 flight sleep cycles. If this proposed Mission Control Simulation Study is funded it will enhance the predictability of Dr. Gundel's computer simulation model by adding an extensive cognitive performance data set.

SUMMARY

The USAF Armstrong Laboratory, Crew Technology Division and the NASA Life Sciences Project Division have

combined their resources to assess the effects of microgravity on Mission Specialists. A subset of performance tests were selected from the Unified Tri-service Cognitive Performance Battery to be implemented on the International Microgravity Laboratory onboard the Space Shuttle. The flight experiment and supporting ground-based studies to determine the reliability and stability of the performance measures were explained. The paper concluded with a suggested control study necessary to differentiate the effects of fatigue and changing work/rest shift cycles from microgravity.

REFERENCES

AGARD (1989). Human Performance Assessment Methods. AGARDograph No. 308. NATO Advisory Group for Aerospace Research and Development.

Benson, A.J. and Gedye, J.L. (1963). Logical processes in the resolution of orienting conflict. RAF Rpt. 259, Farnborough, UK, Royal Air Force Institute of Aviation Medicine.

Bittner, A.C., Jr., Carter, R.C., Kennedy, R.S., Harbeson, M.M. and Krause, M. (1984, September) Performance evaluation tests for environmental research (PETER): evaluation of 112 tasks. NBDL-84R006 (NTIS No. AD152317), New Orleans, LA: Naval Biodynamics Laboratory.

Brainard, G., Rollag, M., Hannon, P., French, J. and Storm, W. (1990). Effects of bright illumination on plasma melatonin in normal volunteers during sustained performance. Proceedings of the Vth Colloquium of the European Pineal Study Group, London, UK.

Brainard, G., Hannon, P., French, J. and Storm, W. (1990). Effects of bright illumination on plasma cortisol during sustained performance. Proceedings Society of Light Treatment of Biological Rhythms, New York.

Carpenter, J.A. and Ross, B.M. (1965). Effect of Alcohol on Short Term Memory, Quarterly Journal of Studies on Alcohol, 26, 561-579.

Carter, R.C., Krause, M., and Harbeson, M.M. (1986). Beware of the reliability of slope scores for individuals. Human Factors, Vol. 28, 673-683.

Carter, R.C., and Wolstad, J.C. (1985). Repeated measurements of spatial ability with the Manikan test. Human Factors, 23, 587-591.

Christensen, J. and Talbot, J. (1986). Review of Psychological Aspects of Space Flight. Aviation Space and Environmental Medicine, Vol 57, 203-212.

Covault, C. (1988, January). Record Soviet manned space flight raises human endurance questions. Aviation Week and Space Technology, p. 25.

Damos, D.L., Bittner, A.C., Jr., Kennedy, R.S., Harbeson, M.M. and Krause, M. (1984). Performance evaluation tests for environmental research (PETER): Critical tracking task. Perceptual and Motor Skills, 58, 567-573.

Damos, D.L. and Lyall, E.A. (1984). The effect of asymmetric transfer on dual-task assessment of voice technology. Proceedings of the 28th Annual Meeting of the Human Factors Society, Santa Monica, CA: Human Factors Society.

Damos, D.L. and Wickens, C.D. (1980). The identification and transfer of timesharing skills. Acta Psychologica, 46, 15-39.

Damos, D.L. (1991). Multiple Task Performance. London, UK, Taylor & Francis, LTD.

Dott, A.B. and McKelvy, R.K. (1977). Influence of Ethyl Alcohol in Moderate Levels on Visual Stimulus Tracking, Human Factors, 19, 191-199.

Fitts, P.M. and Peterson, J.R. (1964). Information capacity of discrete motor responses, Journal of Experimental Psychology, 67(2).

Frazier, M.L. and Crombie, R.B. (1982) Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, AFFTC-TP-82-5, Edwards AFB, CA: Air Force Flight Test Center.

French, J., Hannon, P. and Brainard, G. (1989). Bright lights and fatigue degraded performance. Annual Review of Chronopharmacology, 7, 45-49.

Gevins, A.S., Cutillo, B.A., Fowler-White, R.M., Illes, J. & Bressler, S.L. (1988). Neurophysiological patterns of operational fatigue: preliminary results. NATO/AGARD Conference Proceedings, 432, pp. 22-1 to 22-7.

Gevins, A.S. and Morgan, N.H. Flight helmet EEG system. (1990). USAFSAM-TP-88-7, Brooks AFB, Tx: USAF School of Aerospace Medicine.

Hartzell, E.J., Gopher, D., Hart, S.G., Lee, E. and Dunbar, S., (1983). The Fittsberg Law: The joint impact of memory load and movement difficulty, Proceedings of the Human Factors Society 27th Annual Meeting. (presentation only)

Hegge, F.W., Reeves, D.L., Poole, D.P., and Thorne, D.R. (1985). Unified Tri-Service Cognitive Performance Battery (UTC-PAB). Design and specification of the battery. JWGD3 MILPERF Rpt. 85-1, Ft. Detrick, MD: U.S. Army Medical Research and Development Command.

Hunter, D.R. (1975). Development of an enlisted psychomotor and perceptual test battery. AFHRL-TR-75-60, Brooks AFB, TX: Air Force Human Resources Laboratory.

Klein, R.H. and Jex, H.R. (1975). Effects of Alcohol on a Critical-Tracking Task, Journal of Studies on Alcohol, Vol 36, No. 1.

Levison, W.H., Baron, S. & Kleinman, D.L. (1969). A model for human controller remnant. IEEE Transactions on Man-Machine Systems, 10, 101-108.

McRuer, D.T. and Jex. H.R. (1967). A review of quasi-linear pilot models. IEEE Transactions on Human Factors in Electronics, 8, 231-249.

Miller, J.C., Takamoto, G.M., Bartel, G.M., & Brown, M.D. (1985). Psychophysiological correlates of long-term attention to complex tasks. Behavior Research Methods, Instruments, & Computers, 17(2), 186-190.

Mohs, R.C., Tinklenberg, J.R., Roth, W.Y., and Kopell, B.S. (1980). Sensitivity of some Human Cognitive Functions to Effects of Methamphetamine and Secobarbital, Drug and Alcohol Dependence, 5, 145-150.

Morris, T. L. (1984). Electrooculographic indices of fatigue-induced decrements in flying related performance. Doctoral Dissertation, College Station, TX: Texas A&M University, conducted at USAFSAM.

Nesthus, T.E., Schiflett, S.G., Bomar Jr., J.B., and Holden, R.D. (1988). The effects of different breathing gas concentrations on task performance at high altitude. Aerospace Medical Association Meeting, New Orleans, LA.

Newman, D.J. and Bussolari, S.R. (1990). Dual-Task performance on an interactive human/computer space shuttle flight experiment. International Space Association Meeting, paper #90-035.

O'Donnell, R.D. and Eggemeier, F.T. (1986). Workload assessment methodology. In K.R. Boff, et.al., (eds), Handbook of Perception and Human Performance, Vol II, New York: Wiley, 42-1 to 42-49.

Perez, W.A., Masline, P.J., Ramsey, F.R., and Urban, K.E., (1987). Unified Tri-Services Cognitive Performance Assessment Battery: Review and Methodology. AAMRL-TR-87, Wright Patterson AFB, Dayton, OH: Armstrong Aerospace Medical Research Laboratory.

504

I

Phillips, W.A (1974). On the distinction between sensory storage and short-term visual memory. Perception & Psychophysics, 16(2), 283-290.

Ratino, D., Repperger, D., Goodyear, C., Potor, G., and Rodriguez, L. (1988). Quantification of reaction time and time perception during space shuttle operations. Aviation, Space, and Environmental Medicine, 59, 220-224.

Reader, D.C., Benel, R.A., & Rahe, A.J. (1981). Evaluation of a manikin psychomotor task. USAFSAM-TR-81-10, Brooks AFB, TX: USAF School of Aerospace Medicine.

Reeves, D.L., Thorne, D.R., Winter, S.L., and Hegge, F.W. (1989). The unified tri-service cognitive performance assessment battery (UTC-PAB): Hardware/software design and specifications. NAMRL SR89-1 (JWGD3 MILPERF SR89-1), Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Ross, H., Schwartz, E., and Emmerson, P. (1987). The nature of sensorimotor adaptation to altered G-levels: Evidence from mass discrimination. Aviation, Space, and Environmental Medicine, 58, 148-152.

Ross, B.M. and Chambers, R.M. (1967). Effects of Transvers G-Stress on Running Memory, Perceptual and Motor Skills, 24, 423-435.

Schiflett, S.G., Linton, P.M., and Spicuzza, R.J. (1980). Inflight Evaluation of a Pilot Workload Assessment Device. AGARD/NATO, No. 312, Impact of New Guidance and Control Systems on Military Aircraft Cockpit Design.

Shingledecker, C.A. (1984). A task battery for applied human performance assessment research. AFAMRL-TR-84-071, Wright-Patterson AFB, Ohio: Air Force Aerospace Medical Research Laboratories.

Stepanova, S. (1975). Biorhythmological screening of cosmonauts. Kosmicheskaya Biologiya I Avia-kosmicheskaya Meditsina, 9, 40-46.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. Acta Psychologica, 30, 276-315.

Storm, W.F., Dowd, P.J., & Boll, P.A. (1990, April). Sleep, Fatigue, and Performance in Rail Garrison Habitability Exercise Phase III. Norton AFB, CA: Ballistic Systems Division, Strategic Air Command.

Thorne, D.R., Gensen, S.G., Sing, H.C., and Hegge, F.W. (1985). The Walter Reed Performance Assessment Battery. Neurobehavioral Toxicology and Teratology, 7, 415-418.

Thomas, J.R. and Schrot, J. (1988) Naval medical research institute performance assessment battery (NAMI PAB). NAMRI-88-7, Washington, D.C.: Naval Aviation Medical Research Institute.

Wickens, C.D., Vidulich, M., Sandry, D.L. & Schiflett, S.G. (1981). Factors Influencing Auditory and Speech Channels in Complex Systems. Proceedings, 25th Annual Meeting of the Human Factors Society, Santa Monica, CA: Human Factors Society.

Wickens, C.D. (1986). The effects of control dynamics on performance. In K.R. Boff, et.al., (eds), Handbook of Perception and Human Performance, Vol II, New York: Wiley, 39-1 to 39-60.

Session H3: MEASUREMENTS, TOOLS, AND ANALYSIS - II

Session Chair: Col. Donald Spoon

N92-22336

Transportable Applications Environment (TAE) Plus
a NASA tool used to develop and manage Graphical User Interfaces

Martha R. Szczur

NASA/Goddard Space Flight Center
Greenbelt, MD 20771 USA
mszczur@postman.gsfc.nasa.gov

ABSTRACT

TAE Plus was built at NASA's Goddard Space Flight Center to support the building of graphical user interfaces (GUIs) for highly interactive applications, such as realtime processing systems and scientific analysis systems. It is a general purpose portable tool that includes a *What You See Is What You Get* (WYSIWYG) WorkBench which allows user interface designers to layout and manipulate windows and interaction objects. The WorkBench includes both user entry objects (e.g., radio buttons, menus) and data-driven objects (e.g.,dials, gauges, stripcharts), which dynamically change based on values of realtime data. This paper discusses what TAE Plus provides, how the implementation has utilized state-of-the-art technologies within graphic workstations, and how it has been used both within and outside NASA.

BACKGROUND

Emergence of graphical user interfaces

With the recent emergence of sophisticated graphic workstations and the subsequent demands for highly interactive systems, designing and developing good user interfaces has become more complex and difficult. Prior to the graphic workstations, the application developer was primarily concerned with developing user interfaces for a single monochrome 80x24 alphanumeric character screen with keyboard user entry. With high resolution bit-mapped workstations, the user interface designer has to be cognizant of multiple window displays, the use of color, graphical objects and icons, and various user selection techniques (e.g., mouse, trackball, tablets).

High resolution graphic workstations also provide system developers with the opportunity to rethink and redesign the user interfaces (UI) of their next generation applications. For instance, in a command and control environment, many processes run simultaneously to monitor a particular operation. With modern graphic workstations, time critical information concerning multiple events can be displayed concurrently on the same screen, organized into different windows in a variety of graphical and textual presentations. As today's workstations inspire more elaborate user interfaces, the applications which utilize their graphics capabilities increase in complexity. Thus, an interactive tool that simplies the process of building user interfaces becomes an important productivity element within an application development environment.

Requirements for a prototype-to-operational development environment

In building a user interface development tool, we wanted to estab-lish an integrated environment that allows easy prototyping of an application user interfaces, and also, provides for a smooth transition from the prototyped system into the base operational application. This environment would satisfy the following objectives:

- separate the user interface from the application,
- provide tools to allow interactive design/change/save of user interface elements,
- take advantage of the latest hardware technology,
- support rapid prototyping,
- manage the user interface,
- develop tools for increasing application development productivity,
- provide the application with high level runtime services, and
- allow portability to different computing environments.

WHAT DOES TAE PLUS PROVIDE?

To meet the defined goals, services and tools were developed for creating and managing window-oriented user interfaces. It became apparent, due to the flexibility and complexity of graphical user interfaces, that the design of the user interface should be considered a separate activity from the application program design. The interface designer can then incorporate human factors and graphic art techniques into the user interface design. The application programmer needs only to be concerned about what results are returned by the user interaction and not the look of the user interface.

In support of the user interface designer, an interactive *WorkBench* application was implemented for manipulating interaction objects ranging from simple buttons to complex multi-object panels. As
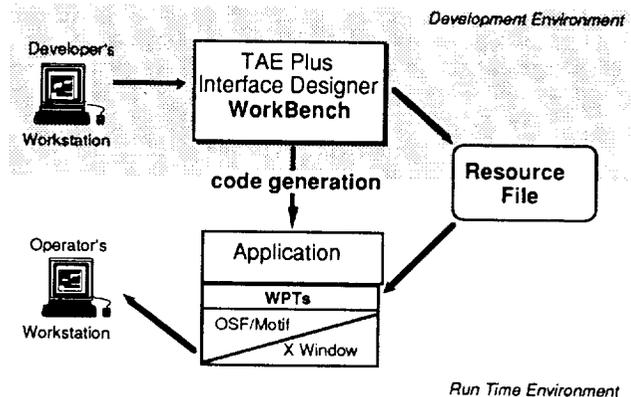
Figure 1. TAE Structure

illustrated in Figure 1, after designing the screen display, the Work-Bench saves the specification of the user interface in resource files, which can then be accessed by application programmers through a set of runtime services, Window Programming Tools (WPTs). Guided by the information in the resource files, the routines handle all user interactions. The WPTs utilize standard underlying software (e.g., MIT X Window System™, Open Software Foundation's Motif™) to communicate with the graphic workstations. [Ref. 2,3] As a further aid to the UI developer, the WorkBench provides an option to generate the source code which will display and manage the designed user interface. This gives the programmer a working template into which application-specific code can be added.

INTERACTION OBJECTS AS BUILDING BLOCKS

The basic building blocks for developing an application's graphical user interface are a set of interaction objects. All visually distinct elements of a display that are created and managed using TAE Plus are considered to be interaction objects and they fall into three categories: user-entry objects, information objects, and data-driven objects. *User-entry objects* are mechanisms by which an application can acquire information and directives from the end user. They include radio buttons, check boxes, text entry fields, sliding scaler, scrolling text lists, pulldown menus and push buttons. *Information objects* are used by an application to instruct or notify the user, such as contextual on-line help information displayed in a scrollable


Figure 2. TAE Plus User Interface Interaction Objects

static text object or brief status error messages displayed in a bother box. *Data-driven objects* are vector-drawn graphic objects which are linked to an application data variable; elements of their view change as the data values change. Examples are dials, thermometers, and strip charts. When creating user dialogues, these objects are grouped and arranged within *panels* (i.e., windows) in the WorkBench.

The use of interaction objects offers the application designer/programmer a number of benefits with the expected payoff of an increase in programmer productivity. The interaction objects provide a consistent look and feel for the application's user interface, which translates into reduced end-user training time, more attractive screens, and an application which is easier to use. Another key benefit is that since the interaction objects have been thoroughly tested and debugged, the programmer is able to spend more time testing the application and less time verifying that the user interface behaves correctly. This is particularly important considering the complexity of some of the objects, and the programming effort it would take to code them from scratch. Refer to Figure 2 for a sample of the TAE Plus interaction objects.

TAE PLUS WORKBENCH

The WorkBench provides an intuitive environment for defining, testing, and communicating the look and feel of an application system. Functionally, the WorkBench allows an application designer to dynamically lay out an application screen, defining its static and dynamic areas. The tool provides the designer with a choice of pre-designed interaction objects and allows for tailoring, combining and rearranging of the objects. To begin the session, the designer needs to create the base panel (i.e., window) into which interaction objects will be specified. The designer specifies presentation information, such as the title, font, color, and optional on-line help for the panel being created. The designer defines both the presentation information and the context information of all interaction items to reside in the panel by using the item specification window (refer to Figure 3). As the UI designer moves, resizes, and alters any of the item's attributes, the changes are dynamically reflected on the display screen.

The designer also has the option of retrieving *palettes* of previously created items. The ability to reuse interaction objects saves programming time, facilitates experimenting with different combinations of items in the prototyping process, and contributes to standardization of the application's look and feel. If an application system manager wanted to ensure consistency and uniformity across an entire application's UI, all developers could be instructed to use only items from the application's palette of common items.

When creating a data-driven object, the designer goes through a similar process by setting the associated attributes (e.g., color thresholds, maximum, minimum, update delta) in the specification panels. To create the associated graphics drawing, the WorkBench provides a drawing tool within which the static background and dynamic foreground of a data-driven object can be drawn, edited, and saved. Figure 4 shows the drawing tool being used to create a *stretcher* data-driven object.

Most often an application's UI will be made up of a number of related panels, sequenced in a meaningful fashion. Through the WorkBench, the designer defines the interface *connections*. These links determine what happens when the user selects a button or a menu entry. The designer attaches *events* to interaction items and thereby designates what panel appears and/or what program executes when an event is triggered. Events are triggered by user-controlled I/O peripherals (e.g., point and click devices or keyboard input).
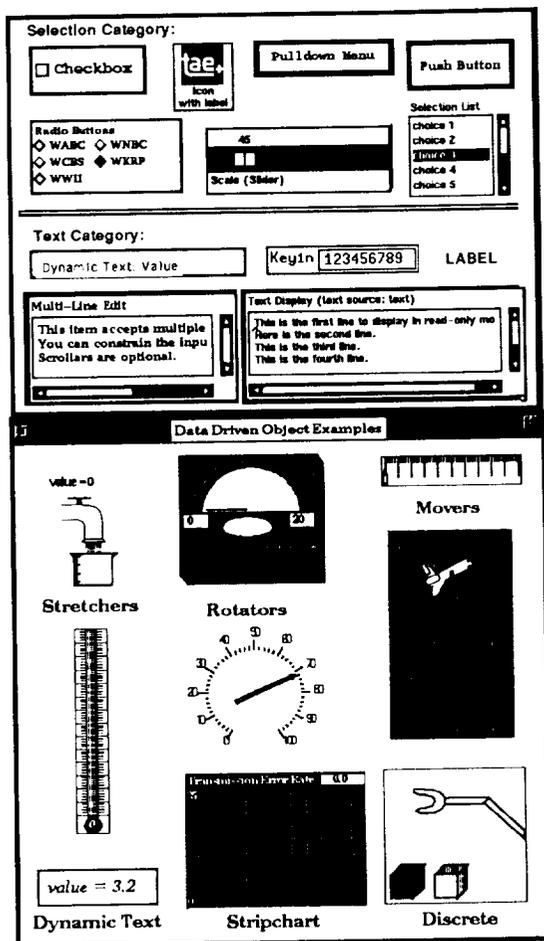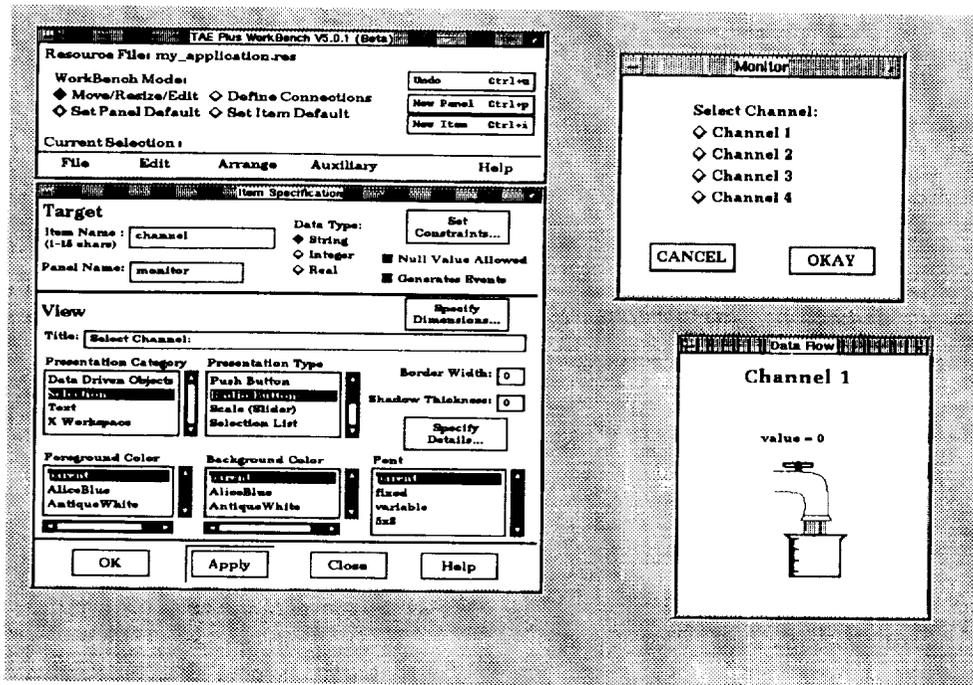
Figure 3. Building a user interface with the WorkBench
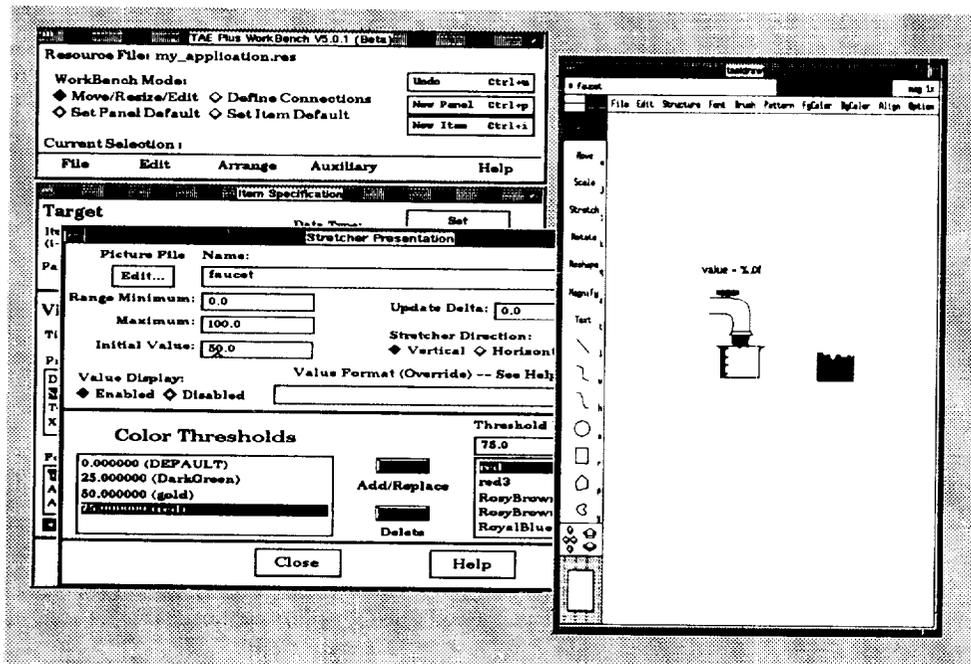
Figure 4. Creating a stretcher data-driven object

TAE Plus also offers an optional help feature which provides a consistent mechanism for supplying application-specific information about a panel and any interaction items within the panel. In a typical session, the designer elects to edit a help file after all the panel items have been designed. Clicking on the edit help option in the Panel Specification Panel brings up a text editor window in which the appropriate information can be entered. The designer can then define any button item or icon item to be the help item for the panel. During the application operation, when the end-user clicks on the defined help item, the cursor changes to a question mark symbol (?). The end-user then clicks on the panel itself or any item in the panel to bring up a help panel containing the associated help text.

Having designed the layout of panels and their attendant items and having threaded the panel and items according to their interaction

510

scenario, the designer is able to preview (i.e., rehearse) the interface's operation from the WorkBench. With this potential to test drive an interface, to make changes, and to test again, iterative design becomes part of the prototyping process. With the rehearsal feature, the designer can evaluate and refine both the functionality and the aesthetics of a proposed interface. After the rehearsal, control is returned to wherever the designer left off in the WorkBench and the designer can either continue with the design process or save the defined UI in a resource file.

Developing software with sophisticated user interfaces is a complex process, mandating the support of varied talents, including human factors experts and application program specialists. Once the UI designer (who may have limited experience with actual code development) has finished the UI, he/she can turn the saved UI resource file over to an experienced programmer. As a further aid to the application programmer, the WorkBench has a "generate" feature, which produces a fully annotated and operational body of code which will display and manage the entire WorkBench-designed UI. Currently, source code generation of C, Ada, and TCL are supported, with bindings for C++ expected in a future release of TAE Plus. The programmer can now add additional code to this template and make a fully functional application. Providing these code stubs helps in establishing uniform programming method and style across large applications or within a family of interrelated software applications.

WINDOW PROGRAMMING TOOLS (WPTs)

The Window Programming Tools (WPTs) are a package of application program callable subroutines used to control an application's user interface. Using these routines, applications can define, display, receive information from, update and/or delete TAE Plus panels and interaction objects. WPTs support a modeless user interface, meaning a user can interact with one of a number of interaction objects within any one of a number of displayed panels. In contrast to sequential mode-oriented programming, modeless programming accepts, at any instance, a number of user inputs, or *events*. Because these multiple events must be handled by the application program, event-driven programming can be more complex than traditional programming. The WorkBench's auto-generation of the WPT event loop reduces the risk of programmer error within the UI portion of an application's implementation.

As mentioned earlier, the WPT package utilizes the MIT X Window System as its base windowing system. One of the strengths of X is the concept of providing a low-level abstraction of windowing support (Xlib), which becomes the base standard, and a high-level abstraction (X toolkits), which has a set of interaction objects (called "widgets" in the X world) that define elements of a UI's look and feel. The current version of TAE Plus (V5.1) operates with the X11R4 and uses the OSF/Motif toolkit, widgets and window manager.

The WPTs also provide a buffer between the application program and the X Window System services. For instance, to display a WorkBench-designed panel, an application makes a single call to Wpt_NewPanel (using the *panel name* specified in the WorkBench). This single call translates into a function that can make as many as 50 calls to X Window System routines. For the majority of applications, the WPT services and objects supported by the WorkBench provide the necessary user interface tools and save the programmer from having to learn the complexities of programming directly with X. This can be a significant advantage, especially when considering the learning curve differential between 26 WPT routines versus over 400 X Toolkit intrinsics and over 200 Xlib services. Refer to Figure 5 for a sample list of the WPTs.

IMPLEMENTATION

The TAE Plus architecture is based on a separation of the user interaction management from the application-specific software. The current implementation is a result of having gone through several prototyped and beta versions of a WorkBench and user interface support services during the 1986-89 period, as well as building on several data structures from an earlier alphanumeric-oriented UI management system, TAE *Classic*.

The "Classic" portion of the TAE Plus code (≈ 60,000 LOC) is implemented in the C programming language. In selecting a language for the WorkBench and the WPT runtime services, we felt a "true" object-oriented language would provide us with the optimum environment for implementing the TAE Plus graphical user interface capabilities. (See Chapter 9 of Cox [Ref. 4] for a discussion on the suitability of object-oriented languages for graphical user interfaces.) We selected C++ [Ref. 5] as our implementation language for several reasons [Ref. 6]. For one, C++ is becoming increasingly popular within the object-oriented programming community. Another strong argument for using C++ was the availability of existing, public domain, X-based object class libraries. Utilizing an existing object library is not only a cost saver, but also serves as a learning tool, both for object-oriented programming and for C++. Delivered with the X Window System is the *InterViews* C++ class library and a drawing utility, *idraw*, both of which were developed at Stanford University. [Ref. 7] The *idraw* utility is a sophisticated direct manipulation C++ application, which we integrated into the WorkBench to support creating, editing and saving the graphical data-driven interaction objects.

AVAILABILITY AND USER SUPPORT

After two years of prototyping and developing beta versions of the TAE Plus, an *industrial strength* version of TAE Plus (Version 4.1) was released in February 1990. A year later, in April 1991, the latest version, TAE Plus V5.1, was released. It is available for public distribution, at a minimal license fee, from the Center of Software Management and Information Center (COSMIC), a NASA distribution center. While TAE Plus base development and testing is done on a Sun workstation under UNIX within the R&D laboratory at GSFC, TAE Plus is also ported onto a variety of UNIX workstations (e.g., Apollo, Vaxstation II, Decstation 3100, HP9000, and Macintosh II with A/UX.) TAE Plus is also available and validated on the Vaxstation II under VMS and DECWindows. Other user sites have successfully installed TAE Plus onto the Masscomp, Silicon Graphics Iris and other Unix-based graphic workstations. During the summer of 1991, as Motif 1.1 becomes available on different platforms, TAE Plus will be ported to a variety of other machines.

Wpt_AddEvent	Add other sources for input/output/exception
Wpt_BeginWait	Display busy indicator cursor
Wpt_CloseItems	Close Items on a Panel
Wpt_ConvertName	Get the X Id of a named window
Wpt_Endwait	Stop displaying busy indicator cursor
Wpt_Init	Initializes interface to X Window System
Wpt_ItemWindow	Gets the window Id of the window containing a parameter
Wpt_MissingVal	Indicates if any values are missing
Wpt_New Panel	Displays a user interface panel
Wpt_NextEvent	Gets next panel-related event
Wpt_PanelErase	Erases the displayed panel from the screen
Wpt_PanelMessage	Displays message in "Bother Box"
Wpt_PanelReset	Resets object values to Initial values
Wpt_PanelTopWindow	Gets panel's parent shell window Id
Wpt_PanelWidgetId	Return the Widget Id of a Wpt Panel Widget
Wpt_PanelWindow	Returns the X Id of a panel
Wpt_ParmReject	Generates a rejection message for a given value
Wpt_ParmUpdate	Updates the displayed values of an object
Wpt_Pending	Check if a WptEvent is pending from X, Parm or file.
Wpt_RemoveEvent	Remove a previously registered event
Wpt_SetTimeOut	Set/Cancel timeout for gathering Wpt events.
Wpt_ViewUpdate	Updates the view of a parameter on a displayed panel

Figure 5. The Window Programming Tools (WPTs)

Since the first release of TAE Classic in 1981, we have provided user support through a fully staffed TAE Support Office (TSO). This service has been one of the primary reasons for the success of TAE. Through the TSO, users receive answers to technical questions, report problems, and make suggestions for improvements. In turn, the TSO keeps users up-to-date on new releases, publishes a newsletter, and sponsors user workshops and conferences. This exchange of information enables the Project Office to keep the TAE software and documentation up-to-date and, perhaps most importantly, take advantage of user feedback to help direct future development.

APPLICATIONS USING TAE PLUS

Since 1982 over 800 sites have installed TAE Classic and/or TAE Plus. The applications built or being built with TAE perform a variety of different functions. TAE Classic usage was primarily used for building and managing large scientific data analysis and data base systems (e.g., NASA's Land Analysis System (LAS), Atmospheric and Oceanographic Information Processing System (AOIPS), and JPL's Multimission Image Processing Laboratory (MIPL) system.) Within the NASA community, TAE Plus is also used for scientific analysis applications, but the heaviest concentration of user applications has shifted to support of realtime control and processing applications. This includes supporting satellite data capture and processing, monitor and control of spacecraft and science instruments, prototyping user interface of the Space Station Freedom crew workstations and supporting diagnostic display windows for realtime control systems in ground operations. For these types of applications, TAE Plus is principally used to design and manage the user interface, which is made up of a combination of user entry and data-driven interaction objects. TAE Plus becomes a part of the development life cycle as projects use TAE Plus to prototype the initial user interface design and have this designed user interface evolve into the operational UI.

Outside the NASA community, TAE Plus is being used by an assortment of other government agencies (22%), universities (15%), and private industries (35%). Within the government sector, users range from the National Center for Atmospheric Research, National Oceanographic and Atmospheric Adminstration, U.S. Geological and EROS Data Center, who are developing scientific analysis, image mapping and data distribution systems, to numerous Department of Defense laboratories, who are building command-and-control-related systems. Universities represented among the TAE community include CalTech, Cornell, Georgia Tech, MIT, Stanford, University of Maryland and University of Colorado. Applications being developed by University of Colorado include the Operations and Science Instrument Support System (OASIS), which monitors and controls spacecraft and science instruments and a robotics testbed for research into the problems of construction and assembly in space. [Ref. 8] Private industry has been a large consumer of the TAE technology and a sample of the companies that have received TAE Plus include Apple Computer Inc., Ford Aerospace, Martin Marietta, Computer Sciences Corp., TRW, Lockheed, IBM, Northern Telecom, Mitre Corp., General Dynamics and GTE Government Systems. These companies are using TAE Plus for an assortment of applications, ranging from a front-end for a corporate database to advanced network control center. Northern Telecom, Inc. used TAE Plus to develop a technical assistance service application which enables users to easily access a variety of applications residing on a network of heterogeneous host computers. [Ref. 9] Because of the high cost associated with programming and software-development, more and more software development groups are looking for easy-to-use productivity tools, and TAE Plus is becoming recognized as a viable tool for developing an application's user interface.

NEXT STEPS

The current TAE Plus provides a useful tool within the user interface development environment -- from the initial design phases of a highly interactive prototype to the fully operational application package. However, there are many enhancements and new capabilities that will be added to TAE Plus in future releases.

In the near term, the emphasis will be on enhancement features and upgrades, such as adding the full set of Motif objects and C++ code generation. All the requested enhancements are user-driven, based on actual experience using TAE Plus, or requirement-driven based on an application's design. For example, on the enhancements list are extensions to the connections mechanism, support for importing foreign graphics, and automating the creation and integration of new interaction objects into the WorkBench.

Future advancements include expanding the scope of the Transportable Applications Environment (TAE) to include new tools or technologies. For instance, the introduction of hypermedia technology, 3-D support and the integration of expert system technology to aid in making user interface design decisions are targeted for investigation and prototyping.

CONCLUSION

With the emergence of sophisticated graphic workstations and the subsequent demands for highly interactive systems, the user interface becomes more complex and includes multiple window displays, the use of color, graphical objects and icons, and various selection techniques. Prototyping of different user interface designs, thus, becomes an increasingly important method for stabilizing concepts and requirements for an application. At GSFC, the TAE Plus development team had the requirement to provide a tool for prototyping a visual representation of a user interface, as well as to establish an integrated development environment that allows prototyped user interfaces to evolve into operational applications. TAE Plus is fulfilling this role by providing a usable, generalized, portable and maintainable package of development tools.

TAE Plus is an evolving system, and its development will continue to be guided by user-defined requirements. To date, each phase of TAE Plus's evolution has taken into account advances in virtual operating systems, human factors research, command language design, standardization efforts and software portability. With TAE Plus's flexibility and functionality, it can contribute both more advances and more standardization in user interface development system technology.

ACKNOWLEDGEMENTS

REFERENCES

1. Perkins, D.C., Howell, D.R., Szczur, M.R., "The Transportable Applications Executive -- an interactive design-to- production development system," *Digital Image Processing In Remote Sensing*, edited by J-P Muller, Taylor & Francis Publishers, London, 1988.

2. Scheifler, Robert W., Gettys, Jim., "The X Window System," MIT Laboratory for Computer Science, Cambridge, MA, October 1986.

3. Open Software Foundation, Inc., *OSF/Motif™ Programmer's Reference Manual*, Revision 1.1, 1990

4. Cox, Brad J., *Object Oriented Programming, An Evolutionary Approach*, Addison-Wesley Publishing Company, Reading, MA, 1986.

5. Stroustrup, Bjarne, *The C++ Programming Language*, Addison-Wesley Publishing Company, Reading, MA, 1987.

6. Szczur, Martha R., Miller, Philip, "Transportable Applications Environment (TAE) Plus: Experiences in 'Object'ively Modernizing a User Interface Environment," Proceedings of the OOPSLA Conference, September 1988.

7. Linton, Mark A., Vlissides, John M., Calder, Paul R., "Composing User Interfaces with Interviews," *IEEE Computer*, February, 1989.

8. Klemp, Marjorie, "TAE Plus in a Command and Control Environment", Proceedings of theTAE Eighth Users' Conference, June, 1990

9. Sharma, Alok, et al., "The TAS Workcenter: An Application Created with TAE", Proceedings of the TAE Eighth Users' Conference, June, 1990

COMPUTER AIDED SYSTEMS HUMAN ENGINEERING:
A HYPERMEDIA TOOL

Kenneth R. Boff
Donald L. Monk
Human Engineering Division
Armstrong Laboratory
Wright-Patterson Air Force Base, OH

William J. Cody
Search Technology, Inc.
Atlanta, GA

ABSTRACT

The Computer Aided Systems Human Engineering (CASHE) system, Version 1.0, is a multi-media ergonomics database on CD-ROM for the Apple Macintosh II computer, being developed for use by human-system designers, educators and researchers. Co-developed by agencies of the US Government and industry* it will initially be available on CD-ROM and will allow users to access ergonomics data and models stored electronically as text, graphics, and audio. The **CASHE** *CD-ROM, Version 1.0 will contain the Boff & Lincoln (1988)* **Engineering Data Compendium, MIL-STD-1472D** *and a unique, interactive simulation capability: the* **Perception & Performance Prototyper.** *Its features also include a specialized data retrieval, scaling, and analysis capability and the state-of-the-art in information retrieval, browsing, and navigation.*

INTRODUCTION

Background

The design of effective human system interfaces requires consideration of the variables influencing the operator's ability to acquire and process task-critical information. The basic research literature contains a large body of relevant data regarding human perceptual and performance capabilities and limitations. Too frequently, however, designs have failed to capitalize on the skills of the operator or have made unreasonable demands on the user. The causes are twofold: first, the sheer volume of existing data makes it hard for designers to review or keep abreast of all the relevant literature; and, second, the form in which these data appear makes it difficult for designers to access and interpret them (Boff, 1990).

Over the past decade, a multi-agency US Government supported effort, the Integrated Perceptual Information for Designers (IPID) Project has been underway to aid the accessibility and use of ergonomics data in system design (Boff, 1987a,b,c; Rouse & Boff, 1987; Rouse, Cody, & Boff, 1990). It is formulated around information management objectives geared toward: (1) Identifying, collecting and consolidating human performance data of potential value to human-system design; (2) "Human engineering" the representation and presentation of these data to enable their effective use by system designers (Lincoln & Boff, 1988); (3) Sponsoring of training opportunities to sensitize system designers to the value and application of ergonomics data in the design of human-systems; and (4) Defining and evaluating of integrated media options for aiding system designers in the access, interpretation and application of ergonomics data. Taken together, these efforts at understanding and remediating problems in the transition of ergonomic research to applications have coalesced into a new model -- CASHE -- for the communication of scientific information to practitioners, educators, and researchers.

CASHE System Concept

The long-range goal of Computer Aided Systems Human Engineering (CASHE) is to integrate ergonomics into human-system design decision making. Achieving this objective will require innovative methods and media for enhancing the accession, assimilation, and application of ergonomics that is consistent with accepted system design practice. Over the next five years, CASHE will focus on conducting exploratory R&D to provide a technology basis and to lower technology risks that we anticipate in the full-scale development of a CASHE environment. During this R&D phase, the project is documenting and disseminating its research efforts while producing, in parallel, an evolving series of commercial quality software products delivered on CD-ROM and other advanced media.

---

* Armstrong Laboratory (Wright-Patterson AFB, OH), US Federal Aviation Administration (Washington, DC), US Army Human Engineering Laboratory (Aberdeen Proving Grounds, MD), Naval Ocean Systems Center (San Diego, CA), the Air Force Office of Scientific Research (Bolling AFB, DC), and Search Technology, Inc. (Atlanta, GA).

CASHE Version 1.0 for the Apple Macintosh computer is the first of these products and is described in detail in this paper. In the development of this product, we have explored a range of issues concerned with:

- Conversion of text to hypertext;
- Alternative representations of complex graphics, tables, and illustrations in small desktop displays;
- Information transformation, analysis, and correlation techniques;
- Aids to help users ask "correct" questions of the system;
- Strategies for multi-document portrayal and integration;
- Incorporation of pedagogical aids for system users;
- Navigation in large-scale hypermedia.

Subsequent versions of the CASHE system will migrate to CAD/CAE work station platforms and incorporate advanced capabilities allowing human-system designers in the future to match equipment and workplace characteristics to human behavioral and anthropometric capabilities and limitations.

CASHE VERSION 1.0 SYSTEM DESCRIPTION

Configuration

CASHE, version 1.0, is under development, scheduled for a late 1992 delivery. The target computer platform is the Macintosh II family of computers equipped with, at least, a 13" monochrome monitor and a CD-ROM drive. A hard disk is recommended for storage of user-defined files. Memory requirements have not yet been determined, although it is expected that at least 2 Mb of RAM will be required, 1 Mb for the application program and 1 Mb for the system and finder. The display interface is being developed to accommodate screen sizes of 13" diagonal and larger. Users will be able to manipulate and move windows on the desktop as in other Macintosh applications. This system (a Mac II with 2 Mb memory, 13" monochrome monitor, and CD-ROM disk drive) will run a minimal CASHE Version 1.0. However, at least two other peripherals will be needed for complete CASHE Version 1.0 operation. A color monitor will be required to view color demonstrations. Some auditory phenomena (e.g., stereo effects) will require use of dual speakers and headphones. Users lacking these optional peripherals will be able to access the entries about color or auditory processes, but will not be able to hear/see the associated phenomena in full detail.

Concept of Operations

Similar to other current Macintosh software, we have adopted a metaphor that our primary users should find familiar and which, therefore, will promote learning and ease-of-use. CASHE Version 1.0 is oriented around a desktop metaphor of the system designer's workplace comprised of four major elements: 1) Bookshelf comprised of the Boff & Lincoln, (1988) *Engineering Data Compendium*, *MIL-STD-1472D*, *the Perception & Performance Prototyper* (P³) and user *Project Files,* 2) file management system, 3) file viewer interface, and 4) visualization tools.

At the most basic level, product functions will be integrated through the use of the desktop. Users will access system functions and information through windows, several of which will be specialized to file type (e.g., data viewing and manipulation, outline viewing, entry viewing, interaction with simulations, etc.). Windows can be opened, sized, and positioned on the desktop to permit easy comparison and data transfer among windows. Users will be able to control the number, size, location, and spatial relations (e.g., tiled vs. overlapped) among as many windows as memory limitations will permit, through a window management system similar to other Macintosh applications.

To take advantage of users' knowledge of and operations with the hardcopy media on which the present product is based, we have adopted a book metaphor for the information sources showcased in CASHE Version 1.0. Hence, the *Compendium*, *MIL-STD-1472D*, P³ and user Project Files each emulate books on a shelf. The first three volumes may be opened and closed, leafed through, annotated, tagged with bookmarks, and include a facility for comparing "pages," etc. The fourth volume may only be opened, closed, and leafed through. The many operations that one normally performs with books are retained or augmented, as feasible.

The file management system will be consistent with that used for all Macintosh applications (documents stored within named folders, and folders embedded within folders). In addition, the product supports many functions that are unique to electronic documents including rapid information retrieval through query and search, linking, merging with other documents and software, etc.

The file viewers are the primary user interface to view, interact, manipulate, and analyze the information contained in the document entries. Each viewer provides specific functionality aimed at providing the user easy access to information in a logical and coherent framework. Control of the interface abides by Macintosh guidelines to the extent possible and desirable for our users' needs (Apple Computer, 1987).

Visualization tools are provided to aid the user's understanding and application of the information available within the reference sources on the Bookshelf. First, users are able to manipulate and transform quantitative and graphical relationships contained in the information base or brought in from external sources through a DataViewer interface. Secondly, we have provided the user with a test bench

simulator that we refer to as the "Perception and Performance Prototyper" (P^3). With the aid of a test bench metaphor, P^3 allows the user to manipulate and interactively test experimental variables discussed in references on the Bookshelf. In addition to direct accessibility of P^3 from the desktop, the P^3 test benches will also be available as a browsable reference on the Bookshelf. Finally, we have augmented many existing illustrations, data functions and tables found in the *Compendium* and *MIL-STD-1472D*. Our objective here was to aid the user to understand the illustration or data while keeping faithful to the intent of the original documents. Hence, all original figures and tables are fully available to the user and are augmented by animations and alternative representations that can be "activated" at the user's option. For example, audio illustrations are provided of aural phenomena that exist only as text descriptions in the original sources.

1. Bookshelf

The Bookshelf in Version 1.0 (See Figure 1), is comprised of four volumes:

A. The *Engineering Data Compendium* (Boff and Lincoln, 1988) was developed through the joint efforts of multiple agencies within the Department of Defense, NASA and NATO AGARD as a standardized ergonomics data resource for system designers. In particular, systematic attention was given during the development of the *Compendium* to: (1) defining approaches to communicate human factors data to designers -- that is, to determine appropriate level of technical content and presentation format, style, and terminology; and (2) enhancing the accessibility of specific technical information relevant to design problems -- that is, providing the user with a reliable and simple means for obtaining the specific data needed. It provides reliable information gleaned from over 70 research areas dealing with performance capabilities and limitations of the human operator. Supplemented by over 2,000 figures, tables and illustrations, 1136 individual entries treat parametric data, models, principles, and quantitative and non-quantitative laws.

B. *MIL-STD-1472D* (Department of Defense, 1989) is a military standard for human engineering design criteria for systems, equipment, and facilities. This standard is a legal reference of human engineering design criteria, principles, and practices. The information contained in this standard is typically referenced by government specifications to ensure human-system integration and enhanced system efficiency, reliability, safety, trainability and maintainability in acquired systems. The standard includes over 80 figures and tables.

C. The *Perception & Performance Prototyper* (P^3) is an interactive simulator that enables users to experience and explore behavioral phenomena
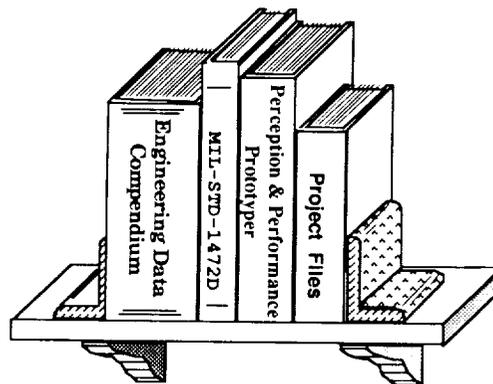
Figure 1. The CASHE Bookshelf

contained in the *Compendium* and *MIL-STD-1472D*. By means of a "test bench" metaphor, users can tailor a simulation to match the conditions of a specific application and immediately experience its effects on their own performance. Collectively, these simulations may be accessed as an independent volume on the Bookshelf through which users may browse or search.

D. The *Project File* is the collection of information which the user may add to customize the system, store files and annotations created during a session, and otherwise augment the information from the other three volumes. Project files may be created, saved, and retrieved in much the same way as documents in word processing applications or data files in database applications. Upon initiating the program, the user will have the option to open a new project file or select from a list of existing project files created in prior interactions.

2. File Management System

The file management system allows the user to access and navigate the CASHE information base found in the four Bookshelf volumes. Access into the information base is accomplished via browsing and search/query results. Users can browse the table of contents, index, design checklists, author references, glossaries, and general access taxonomies. In addition, Boolean logic applied to full text search commands can be used to locate text anywhere in the selected document or documents.

Users may annotate the information base to enhance its personal meaning and value. Annotations will include attaching notes to information objects (notes may contain text and graphics), linking together objects that the user associates, marking objects for subsequent recall, and developing personal indices of terms and then linking these terms to objects in the information base. These annotations are contained in the fourth volume on the Bookshelf, *Project File*.

516

In addition to accessing, navigating, and annotating the information, users may also access objects (other applications, system services, Macintosh desk accessories, etc.) external to our product. Files from other applications that can be recognized by our software may be imported. Files that are produced in the course of using the product may be saved and/or exported. Users will be able to copy most read-only files (not programs) that reside on the compact disk. They will also be able to perform a full range of editing functions on their own files generated within the application.

3. File Viewer Interface

Our interest in providing users with a uniform interface to the reference sources on the Bookshelf eventually raised serious pragmatic and philosophical questions regarding the extent to which we should port documents intact into the hypermedium versus augmenting their presentation to take full advantage of the hypermedium. Re-authoring documents is a costly and highly uncertain process. On the other hand, as ergonomists, we could see the potential lost opportunities of not capitalizing on the advantages of hypermedia.

Given the R&D basis of this project, we decided to experiment with augmentation of ill-structured documents. The granular level at which reference sources on the CASHE Version 1.0 Bookshelf are accessible to the user is the "entry." It is a self-contained unit of information that is addressable by key terms and cross-referencing. Since the *Compendium* was originally designed as a computer-accessible document, it is already parsed into consistent entry units with a highly uniform substructure (See Table 1) that is amenable to conversion to hypertext. *MIL-STD-1472D* is a hierarchically-structured document without a consistent entry level substructure. With the aid of subject matter experts, we augmented the existing document with an entry-like structure that we believe is naturally implied by its content.

To help the user locate and interpret pertinent information, a standardized presentation format was developed for entries in the *Compendium* which is tailored to the needs of the design engineer. This format, has evolved over several years through an iterative process of review and discussion with the user community, sponsors, and consultants. In hard copy form, it represents our best attempt at "human factoring" the presentation of relevant perceptual and performance data. All of the data contained in the hardcopy edition are included in the electronic edition. Although the physical layout differs from version. Although the physical layout differs from the hardcopy edition, a format was developed to optimize the information requirements of designers within the limited physical screen size. This electronic entry allows us to make use of advanced techniques for data retrieval, browsing, navigation, and display.

Table 1. Modular substructure of an entry from the Engineering Data Compendium

Title - The title provides a concise description of the entry content.

Key Terms - This section lists terms that relate to the topic discussed in the entry. Along with key words in the entry title, these key terms can be used to verify entry content and serve as access terms in an index search for related information.

General Description - In entries presenting basic data, this section summarizes the general findings, conclusions, and trends in the data. For entries presenting perceptual/performance models, laws, or principles, it provides a precise description or definition and indicates the general purpose for which the model, law, or principle was developed.

Applications - This section describes general areas of application for the information in the entry; specific types of displays, control systems, task environments, etc., for which the information might be useful; and, where pertinent, general procedures for application.

Methods - Entries presenting basic data contain a Methods section that describes how the data were collected.

Experimental Results - When an entry reports findings of a research study, it contains an Experimental Results section that provides a more detailed discussion of the data than the General Description. The Experimental Results includes graphics or tabular presentation of the data, an enumeration of the major findings and trends in the data, and an indication of their meaning or significance.

Empirical Validation - This section is found in entries that treat a model, law, or principle. It includes a description of the methods used in empirical tests of the model, law, or principles and reports the results and scope of the validation studies.

Constraints - This section describes features or limitations of the information in the entry that may affect its application; stimulus or subject characteristics, environmental conditions, etc., that may influence the results or effect reported; criteria that must be met for proper application; and limits on the class of response, stimulus, task environment, etc., to which the information can be applied.

Key References - This section provides full bibliographical data for several reference sources that contain more detailed information on the entry topic. The original source of the data, model, etc., presented in the entry is marked with an asterisk. References are listed alphabetically and numbered consecutively.

Cross References - This section cites Compendium entries that treat related topics or provide pertinent background information useful in understanding or interpreting the information in the entry.

Each electronic entry is composed of the following elements:

• Identification region. At the top of each entry is its title, its section, and the grouping within the section where the entry is located. Disorientation is a major problem in hypermedia databases. This on-screen verbal format design, along with hierarchical maps (not shown), are designed to minimize disorientation. By selecting one of the arrows to the left of the entry title, the user may browse the next or previous entry.

• User annotation region. This region allows access to user-created annotation in the form of bookmarks, links, and notes.

• "Brass Plate". The brass plate allows access to the four primary elements of an entry: (1) Figures, (2) Text, (3) Tables, and (4) Prototypers. Each of these four elements has an independent "viewer" associated with it. Designers tend to be data rather than text oriented. Hence, the brass plate icons for figures, text, tables, and prototypers are prominently placed in the EntryViewer to draw user attention. By clicking on these icons, the user can quickly jump to the appropriate graphic or table.

• Element viewer. This viewer displays the selected figure, text, or table in a large window covering about 60% of the screen.

• Options panel. For each element viewer, there is an options panel, customized for each document. Here the user can select the various options associated with each viewer and with that document. Each of these viewers and associated option's panel are described in detail below.

A. TextViewer: TextViewer (Figure 2) displays the text portions of an entry. A large text window is available for displaying substructural components of any of the reference sources on our Bookshelf. Figure 4 exemplifies TextViewer for the substructural components of a typical *Compendium* entry. Users can select among the nine components shown by clicking the desired radio button. Subsequent navigation is accomplished by either reselecting a radio button or by clicking on one of the two arrows below the text window to move to the previous or next component. This data entry format is designed to facilitate quick and efficient viewing of the data entries and navigation to other pertinent entries. From the displayed Cross References component, the user may go to any of the given cross references by simply double clicking on the reference.
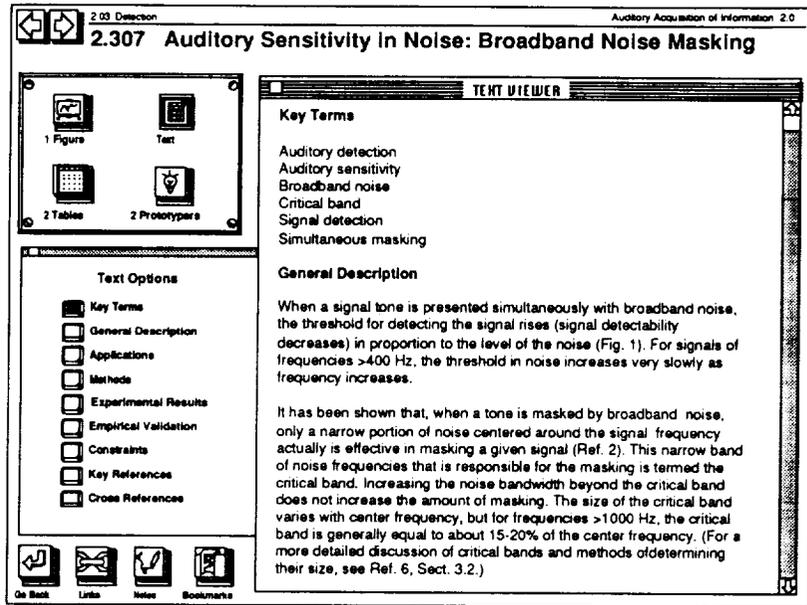
Figure 2. Electronic representation of an entry from the *Engineering Data Compendium* showing TextViewer.

B. FigureViewer: FigureViewer (Figure 3) displays both data graphs and illustrations. Its associated options panel allows the user to select figure panels, control the display of overlays, zoom the figure scaling in or out, and turns the figure caption on or off. If demonstrations are available for the figure, then the user can select the demo button for either an animation or interactive demonstration which further illustrates the entry. If the figure is a data graph, then by pressing the data analysis button, the user can access DataViewer to further explore analytic relationships.

C. TableViewer: TableViewer (Figure 4) displays entry tables. The first row and column of a table contain the row/column titles. If the user scrolls horizontally through the table, then the first column remains fixed with the remaining columns "sliding under" the column titles. Likewise, if the table is scrolled vertically, then the top row of titles remains fixed. The table options panel allows the user to select multiple table panels and turn the table caption on or off.

4. Visualization Tools

Two specialized human perception and performance visualization tools (DataViewer and P³) exist to further aid the user's understanding of the data.

A. DataViewer: DataViewer provides the user with the capability to view, manipulate, and display data. Multiple DataViewer windows may be opened (with only one active), and multiple DataViewer windows may be associated with a single entry.

518
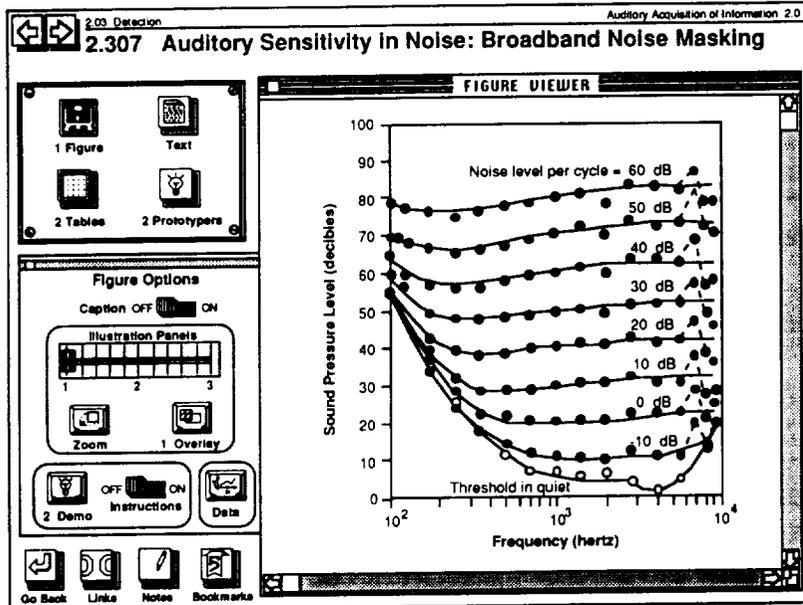
Figure 3. Electronic representation of an entry from the
Engineering Data Compendium **showing FigureViewer.**
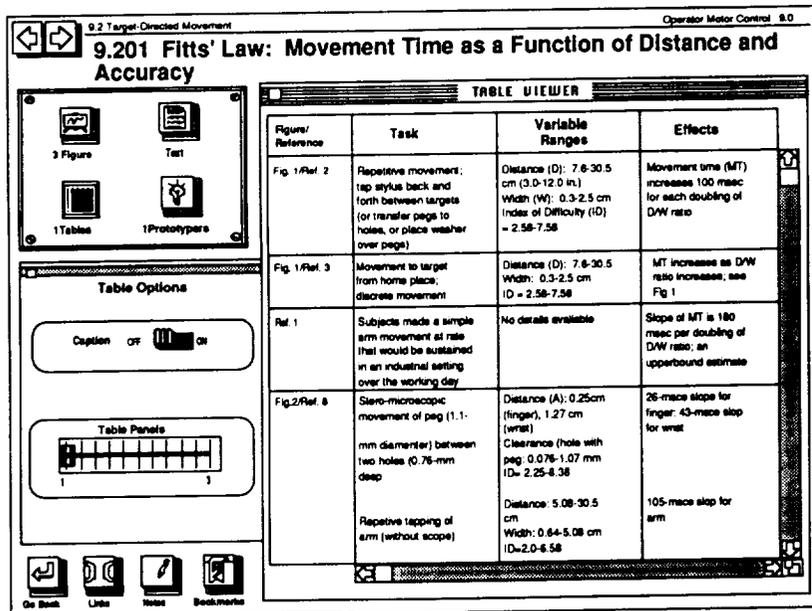
Figure 4. Electronic representation of an entry from the
Engineering Data Compendium **showing TableViewer.**

Figure 5 portrays the selection of an existing XY graph, definition of its axes, digitization and storage of the resulting XY values in a table, and transformation of those data values into a new function, Z. Figure 6 illustrates this concept of data transformation available to the users of "DataViewer."

Four primary functions are available in DataViewer: 1) Data Definition, 2) Data Acquisition, 3) Data Transformation, and 4) Data Presentation. The data transformation function will support the user in exploring a phenomenon's mathematical relationships. Three classes of transformation will be supported: 1) monadic, single variable, operations, 2) dyadic, two variable, operations, and 3) analyses including descriptive statistics and linear/nonlinear polynomial curve fitting.

The data presentation function will support the user in exploring a phenomenon's graphical modes. Five display formats will be offered: 1) histograms, 2) scatterplots, 3) line graphs, 4) pie charts, and (5) tabular. For each of these formats, the user will be able to directly manipulate the graphs to perform such functions as rescale axes, crop or refocus the area of interest, and select/code a point or group of points.

B. Perception & Performance Prototyper (P³): The P³ is a unique feature of Version 1.0 which allows the user to manipulate and experience alternative representations of the technical data found in the reference sources on the Bookshelf. Figure 7 shows an example of how a user may use the Perception and Performance Prototyper:

After consulting data in the *Compendium* concerned with how noise level may affect audio sensitivity, the user may have residual uncertainty regarding the relevance of the data to specific conditions. Using the P³, users will actually be able to access a test bench and custom control panel that will support test and experimentation with the combinations of variables with which they are more directly

2.307 Auditory Sensitivity in Noise: Broadband Noise Masking

DATA VIEWER

Data Points: [#]   X: [6.98] Y: [484.8]

100	55
170	35
225	24
315	20
500	15
700	12
1000	11
1100	10
4000	11

Sound Pressure Level (decibels)

Frequency (hertz)

[Save] [Close]

[Caption] [Frame Axes] [Stop Digitizing] [Data Table] [New] [Cancel]
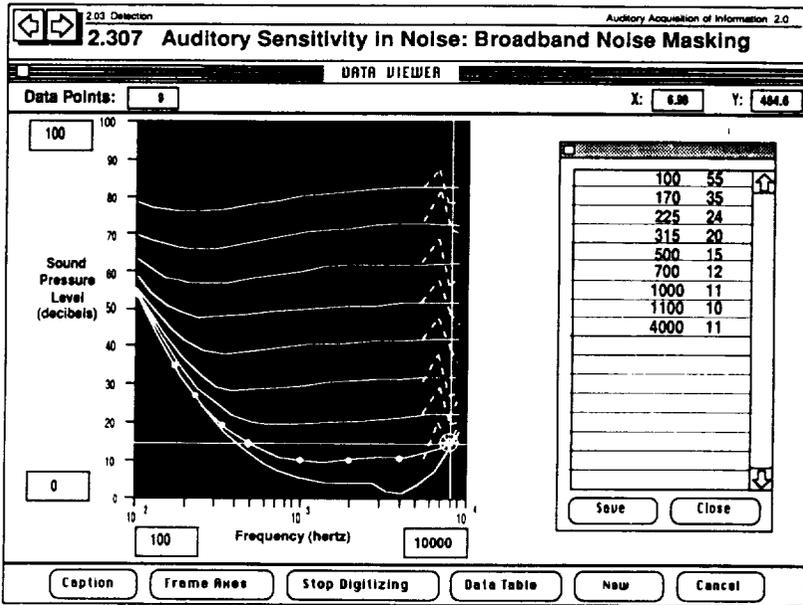
Figure 5. Electronic representation of an entry from the
Engineering Data Compendium **showing DataViewer.**

Data Acquisition

Sound Pressure Level (dB)

60 dB
40 dB
20 dB
0 dB

Frequency (Hz)

Data Transformation

$z=f(x,y,Q)$

X	Y
200	56
400	57
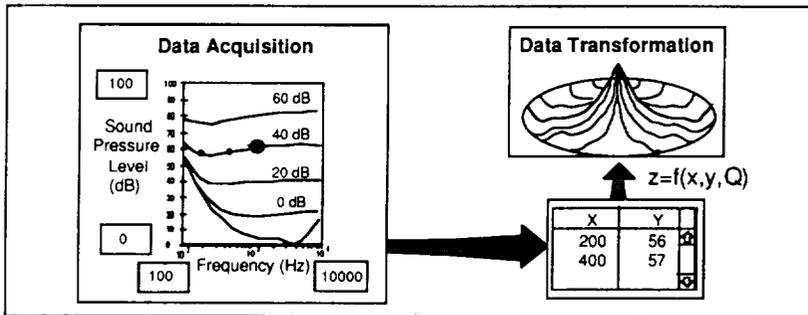
Figure 6. Concept of data transformation available to the users of
"DataViewer."

Control Panel

Audio Masking Simulation

Select Noise
O Soft
● Medium
O Loud

Control Panel

Signal Types
O 44 Hz
● 1000 Hz
O 10000 Hz
O Speech

🔊 Listen

Signal

Noise

Frequency

Distribution
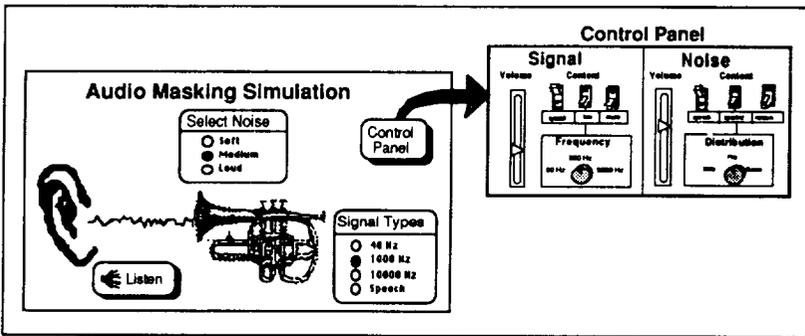
Figure 7. Operation of the *Perception & Performance Prototyper:*
The audio masking test bench.

concerned. In the instance illustrated in Figure 7, the user can dial up relevant signal-to-noise conditions and can then experience the resulting aural phenomenon, first-hand. Minor variations in the phenomenon can be explored by selecting different noise levels or signal types.

If conditional variables can be veridically portrayed, then users will be able to directly experience resultant phenomena. If, on the other hand, conditional variables cannot be faithfully represented, as a result of software or hardware limitations, P[3] will simulate cause and effect relationships with strict adherence to the published data. This approach is employed in CASHE Version 1.0 wherever its availability can be justified by pedagogical criteria. In all instances, users will be able to easily link from and to relevant supporting data found in references on the Bookshelf. In addition to direct accessibility of P[3] from the desktop, the P[3] test benches will also be available as a browsable reference on the Bookshelf.

CONCLUSIONS

In sum, CASHE Version 1.0 offers human-system designers, researchers and educators an unprecedented and unique capability. CASHE supports:
- Rapid and effective access to ergonomics data and models
- Interpretation and application of this information by providing the ability to manipulate and experience alternative representations of these data.

In the future, we will continue to develop CASHE capabilities and implement these through interim commercial products including special purpose software, CD-ROM hypermedia databases, and publications in the open literature which document our ongoing R&D efforts.

REFERENCES

Apple Computer, Inc. (1987). *Human interface guidelines: The Apple desktop interface.* Menlo Park, CA: Addison-Wesley Publishing Company, Inc.

Boff, K. R. (1987a). Matching crew system specifications to human performance capabilities. *Proceedings of the 45th NATO AGARD Guidance and Control Panel Symposium.* Stuttgart, Germany: NATO Advisory Group for Aerospace Research and Development.

Boff, K. R. (1987b). Designing for design effectiveness of complex avionics systems. In *The design, development and testing of complex avionics systems.* Las Vegas, NV: NATO Advisory Group for Aerospace Research and Development. *NATO AGARD CP-417.*

Boff, K. R. (1987c). *The tower of babel revisited: Cross-disciplinary chokepoints in system design.* In Rouse and Boff (Eds.), System design: Behavioral perspectives on designers, tools and organizations (pp. 83-96). New York: North-Holland.

Boff, K. R. (1990). Meeting the challenge: Factors in the design and acquisition of human engineered systems. In H. Booher (Ed.), *People, machines and organizations: A MANPRINT approach to systems integration.* New York: VanNostrand Reinhold.

Boff, K. R. and Lincoln, J. (Eds.) (1988). *Engineering data compendium: Human perception and performance.* (4 Volumes). Wright-Patterson AFB, OH: Armstrong Aerospace Medical Research Laboratory, AAMRL/NATO.

Department of Defense. (1989,14 Mar). *Human engineering design criteria for military systems, equipment and facilities (MIL-STD-1472D).* Washington, DC.

Lincoln, J. E. and Boff, K. R. (1988). Making behavioral data useful for system design applications: Development of the *Engineering Data Compendium. Proceedings of the Human Factors Society 32nd Annual Meeting,* 1021-1025.

Rouse, W. and Boff, K. R. (Eds.) (1987). *System design: Behavioral perspectives on designers, tools and organizations.* New York: North-Holland.

Rouse, W. B., Cody, W. J., and Boff, K. R. (1991). The human factors of system design: Understanding and enhancing the role of human factors engineering. *International Journal of Human Factors in Manufacturing,* 1(1), 87-104.

N92-22338

THE APPLICATION OF INTEGRATED KNOWLEDGE-BASED SYSTEMS FOR THE BIOMEDICAL RISK ASSESSMENT INTELLIGENT NETWORK (BRAIN)

Karin C. Loftin (KRUG Life Sciences, Inc.), Bebe Ly (NASA, JSC), Laurie Webster (NASA-JSC), James Verlander (KRUG Life Sciences, Inc.), Gerald R. Taylor (NASA-JSC), Gary Riley (NASA-JSC), and Chris Culbert (NASA-JSC)

ABSTRACT

One of NASA's goals for long duration space flight is to maintain acceptable levels of crew health, safety and performance. One way of meeting this goal is through BRAIN, an integrated network of both human and computer elements. BRAIN will function as an advisor to mission managers by assessing the risk of inflight biomedical problems and recommending appropriate countermeasures. This paper describes the joint effort among various NASA elements to develop BRAIN and the Infectious Disease Risk Assessment (IDRA) prototype. The implementation of this effort addresses the technological aspects of: (1) knowledge acquisition, (2) integration of IDRA components, (3) use of expert systems to automate the biomedical prediction process, (4) development of a user friendly interface, and (5) integration of the IDRA and ExerCISys systems. Because C Language, CLIPS, (the C Language Integrated Production System), and the X-Window System are portable and easily integrated, they have been chosen as the tools for the initial IDRA prototype.

INTRODUCTION

One of NASA'S goals for long duration space flight is to maintain acceptable levels of crew health, safety and performance. To do this, NASA will monitor crew physiological, psychological and task performance. It also must administer appropriate countermeasures (17,29,37). It is our philosophy that determining the risk of inflight performance problems is the first step of preventing them.

Biomedical risk assessment estimates the probability of a specified human response to a challenge (10,19). The probability estimate is based on epidemiological studies of populations at risk after exposure to extreme conditions. And, certain physiological, psychological and environmental indicators change the estimation of risks for the individual. The specificity of the indicator to predict a human response is the limiting factor in the risk assessment. The acceptable limits of risk have to be understood to manage them.

As the duration of space flight lengthens, the risk and number of biomedical problems will increase. It is crucial to make an assessment and initiate countermeasures. It also is important to predict the impact of the selected countermeasures on crew health, safety and performance. If more than one change in crew status is observed, it is critical to evaluate each countermeasure relative to the others.

A partial solution to the decision-making process required for long duration flights is the application of automated technology. It reduces the volume of data, facilitates data interpretation and resolves incompatible data. For example, expert or knowledge-based systems can automate the diagnostic process that relies on large quantities of related physiological or anatomical data (20,25,40). Each specific discipline has its specific expert system (5,6,7).

Expert systems are commonly rule-based production systems (13). The Software Technology Branch at NASA/JSC has developed a rule-based production system called CLIPS, the C Language Integrated Production Systems (9,12,35). CLIPS and its functions serve as an example of the principles and programming of expert systems (13). CLIPS is being used to automate the prediction process of the Infectious Disease Risk Assessment (IDRA) prototype (see below) and BRAIN (Biomedical Risk Assessment Intelligent Network).

NASA is presently supporting the development of three life sciences expert systems for use on long duration space flight:

1. The IDRA prototype, which assesses the risk of infectious diseases and recommends countermeasures to reduce the risks. The implementation approach and the results of this development are presented in this paper.

2. The Exercise Countermeasures Intelligent System (ExerCISys), which prescribes an exercise protocol to maintain muscle strength and cardiovascular aerobic capacity inflight.

I

3. The Performance Prediction Model (PPM), which assesses and predicts the level of work performance of astronauts. PPM assesses the cognitive sensory motor performance of the individual as it relates to the individual job tasks, or the individual's job as a whole.

These expert systems are independent from each other. They are designed for a single user and the data are not automatically shared. The greatest concern to us is the length of time required for a mission manager to resolve recommendations from these systems and then make real-time decisions.

A solution to this concern is the (BRAIN). The application of knowledge-based systems or artificial intelligence is a vital component of BRAIN. BRAIN is an integrated network for biomedical risk assessment and management. It provides the consensus of multiple experts. We hypothesize that BRAIN will reduce the time required for one to arrive at real-time decisions about biomedical risk analysis and management.

Others outside NASA will benefit from the development of BRAIN. Institutions such as hospitals, medical clinics, boarding schools, military services, and nursing homes for the mentally and physically handicapped are potential users.

IMPLEMENTATION APPROACH

1. Documentation Of Requirements

The preliminary software requirements of BRAIN will be documented according to the IEEE Standards Board and the American National Standards Institute (20).

The BRAIN concept is illustrated as a triangle (Figure 1) with users on the left side of the triangle and expert systems on the right side. The network still permits each user and system to work independently and interact independently with the mission manager. Through BRAIN, each system may access pertinent data from other systems. BRAIN cooperates with the independent expert systems by use of a knowledge base that relates all of them.
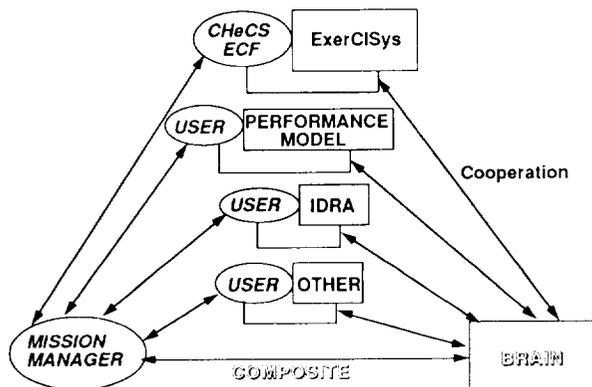
Figure 1. The BRAIN Concept

The functions of BRAIN are to:

• access IDRA, ExerCISys, PPM and other undefined systems for pertinent information.

• assess the biomedical risks and recommend countermeasures.

• function as a clearing house of information to be shared between systems.

• resolve incompatible information given by other expert systems and derive a consensus for a mission manager.

2. System Design
Knowledge Definition

A major activity of this project is to develop the knowledge-based system design. This includes the identification of data sources, knowledge definition, knowledge design and the architecture of the hardware/software environment for BRAIN. The knowledge definition task defines the knowledge requirements of the network and identifies and selects the knowledge sources. The knowledge is acquired, analyzed and extracted. The knowledge design comprises the knowledge representation, i.e., rules, internal fact structure, detailed control structure and preliminary user interface (13).

BRAIN receives input from PPM, IDRA, and ExerCISys, or the user, as illustrated in Figure 1. Other data that BRAIN requires are derived from textbooks, journal articles and reports. Inaccessible data may be simulated as necessary during initial development. The data are stored in a local or working database.

The data structure and network configuration of BRAIN must be compatible with IDRA, ExerCISys and PPM. These expert systems share related data through BRAIN by accessing the working database. We will test the feasibility of IDRA, ExerCISys and PPM to regularly post data that are required by other systems. A standard protocol will be established for each system to access BRAIN and vice versa.

The knowledge base for BRAIN utilizes and interprets the data, predicts the risk of biomedical problems and recommends the appropriate countermeasures. It is retrieved from the sources listed below.

• *Spaceflight Historical Information*

• *Expert Medical and Science Personnel*

• *Texts, Journal Article and Reviews*

• *Epidemiological Studies of Normal Populations*

The resources available in the medical sciences arena and NASA life sciences groups are explored for the knowledge definition of BRAIN. The relationships among IDRA, ExerCISys and PPM are defined by means of workshops, personal consultation and collaboration of existing study results. Experts will be identified so we can model their expertise and to evaluate the demonstration of BRAIN during the developmental stages.

Knowledge Acquisition

Once the knowledge base has been defined for BRAIN, methods will be developed to acquire the specific knowledge. Since a great deal of knowledge has to be acquired for BRAIN, an automated method may be required for that purpose. That method must be consistent and reproducible while extracting information from human experts and written sources.

Knowledge Design

A conceptual design of BRAIN is illustrated in Figure 2. Further definition of the knowledge representation and design is delayed until the Knowledge Acquisition is completed. At that time, more will be known about the structure of the knowledge and how it can best be represented.

It is anticipated that the knowledge may be subjected to a software tool called RuleMaster that uses the ID3 algorithm. The ID3 algorithm analyses empirical data

and derives rules for the knowledge base of BRAIN. Advanced techniques, e.g., CLIPS, will be tested to automate the biomedical prediction process. Other existing and newly developed tools will be evaluated for their best knowledge representation and design capability.

BRAIN will be designed with a learning capability. It will incorporate, by a feed-back mechanism, the experience of an expert . The decisions and interpretations of data obtained from actual test cases are acquired automatically in the knowledge base and new rules are induced. This function is entirely under the control of the appropriate user. But once initiated, it is automatically included in the knowledge base. Tools such as the Automated Reasoning Tool (ART) and the Automated Structured Rule Acquisition (ASTRA) are being used to capture the expertise of exercise physiologists for ExerCISys. ART and ASTRA are being evaluated for application to BRAIN.

Knowledge Verification/Validation

Verification and validation of BRAIN is a vital step throughout the life cycle of its development (9,18). Verification of BRAIN determines that the software is developed according to specifications. Validation determines that BRAIN performs the functions as specified by the requirements and is usable for field testing (11).
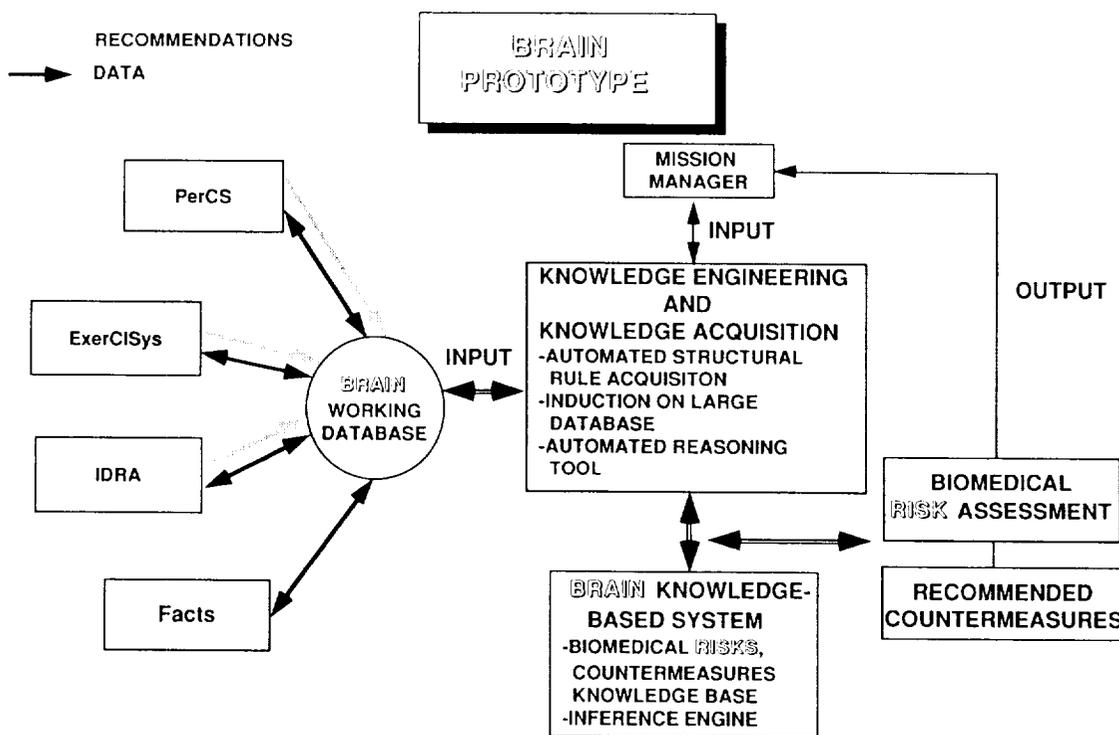

Figure 2. Conceptual Design Of Brain

524

During the knowledge design, verification determines that the design adheres to the requirements. The knowledge base is verified by checking specific details to the level of each rule.

Validation of BRAIN encompasses aspects of:

• determining the validation criteria. (25)

• specifying the sets of input data.

• developing a library of test cases and detailed space flight scenarios.

• validating BRAIN by an independent panel of experts.

• using BRAIN in parallel with the independent operation of PPM, IDRA and ExerCISys and then comparing the results. (11)

After the Preliminary Design Review of the project, the detailed design description will be documented. It will specify the logic and content of the knowledge base, the implementation of the system, hardware requirements, the detailed user interface and the detailed demonstration plan.

The hardware/software environment of BRAIN will be compatible with PPM, IDRA, ExerCISys and Space Station Freedom standards. The development environment that is used to create the software may not run on the identical platform as the demonstration version.

3. User Interface

It is essential for the flight components of BRAIN to have user friendly interfaces. Ease of use is important to whether or not a system is fully utilized. The X-Window System will be used to develop the preliminary user interface. The user interfaces to BRAIN will be designed in accordance with human factors principles and the Space Station Volume of the Man Systems Integration Standards (NASA Std 3000) document. Some of the factors that will be addressed are controls, visual displays and auditory demonstration displays (41). Prior to completion of the final BRAIN design, all interfaces will be empirically evaluated using subjects similar to the typical user. Based upon findings of this study, the design of the interfaces will be refined.

After BRAIN is developed, it will be field-tested during future space flights, bed rest studies, military activities and Antarctic expeditions. New versions of BRAIN will be developed (based on field-test results) to accommodate each test environment. The scientific potential for advancing telescience communications between BRAIN and remote study locations also exists and will be explored. Advanced computer technology promises to assist humans in the 21st century to better cope with the uncertainties of health, safety and performance at home and in the work place.

Table I. Expected results. At the end of the project period the products will function as indicated.

PRODUCTS	FUNCTION
BRAIN SOFTWARE (Interference Machine)	Automates the biomedical risk assessment and prediction process by appropriately executing the rules
	Accesses IDRA, ExerCISys, and PPM
WORKING DATABASE	Contains the facts required for rules
BRAIN KNOWLEDGE BASE	Evaluates and interprets the related outputs of IDRA, ExerCISys, and PPM in a set of rules.
BRAIN OUTPUT	Generates a composite risk analysis and recommendations to assist the user in making real-time decisions
STANDARDIZED PROTOCOL FOR INTEGRATION	A standard procedure to integrate expert systems in BRAIN
COMPUTER HARDWARE CONNECTIONS	Communicate with IDRA, ExerCISys, and PPM
USER-FRIENDLY COMPUTER HARDWARE/SOFTWARE	Increase efficiency, productivity, and quality of the BRAIN output

4. The IDRA Prototype

Because the prevention of infections in space is important, an IDRA prototype was developed (3,28,33). The goal of the functional IDRA prototype is to test the feasibility of using knowledge-based systems for infectious disease risk assessment.

The IDRA prototype focused on respiratory infections, especially influenza. The epidemiology and procedures for preventing, diagnosing and treating influenza are well defined (1,4,38).

Epidemiological studies have evaluated the risk factors and their predictive value for influenza in the general population (8,24,35) and the efficacy of chemotherapeutic prophylaxis (14). Earlier studies investigated the outbreak of influenza in isolated populations, e.g., on aircraft (27), ships at sea (32) and college campuses (23,34). From these sources, we concluded that sufficient information was available to construct a knowledge base about influenza.

Studies show that exercise has a profound effect on the immune system (22,30,31) and sometimes induces changes similar to those arising from the stress of space flight (16). The exercise regimen and related physiological data are factors that must be taken into consideration for the risk assessment of infectious diseases and for prescribing an exercise program. This was exemplified on the MIR when Cosmonaut Gennady Strekalov caught a cold following exercise (reported by the Associated Press, October 18, 1990). The IDRA prototype will be compatible with the ExerCISys prototype. We will integrate IDRA with the ExerCISys as a model for BRAIN.

IDRA Results

The knowledge for the IDRA knowledge base was extracted and analyzed from textbooks and journal articles. We identified the critical indicators that predict the probability of respiratory infections. These indicators were best understood for influenza.

The risk of influenza for an individual is described by general population statistics. It is dependent on an individual's location, age group and level of immunity. This information is encoded in a set of 23 rules using CLIPS. The integrated knowledge-base of IDRA and ExerCISys will contain information about the relationship between exercise, the immune system and infections.

Figure 3 illustrates the major components of the IDRA prototype. A C-based data management tool interacts with all the components of the system. It processes information from the database and from the user interface. The expert system using CLIPS assesses the probability of influenza. It retrieves the information from the data management tool and outputs it to the user interface. A screen displays the probability of infection and illness in the form of a text and a graph.

For the preliminary user interface, we used the X-Window System. All of the tools are portable and compatible with Space Station Freedom requirements.

CONCLUSIONS

• BRAIN can help solve the problems of assessing biomedical risks and performance decrements of humans working in microgravity.

• BRAIN can provide a consensus to the mission manager by surveying independent expert systems.

• The functional IDRA prototype demonstrates that risk analysis for influenza can be automated using C Language, CLIPS, and the X-Window System. The IDRA prototype will be integrated with ExerCISys and used to develop BRAIN.
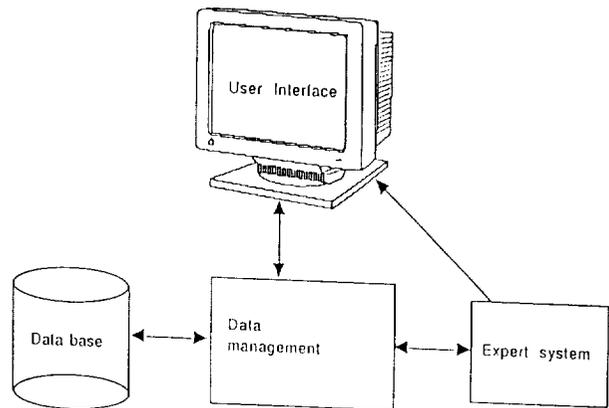
Figure 3. The Major Components Of The IDRA Prototype

REFERENCES

1. Atmar RL, Greenberg SB, Quarles JM, Wilson SM, Tyler B, Feldman S, Couch RB. Safety and pharmacokinetics of rimantadine small-particle aerosol. *Antimicrobial Agents and Chemotherapy* 1990; 34(11):2228-2233.

2. Beall JH, chairman. *IEEE Guide to Software Requirements Specifications.* New York: The Institute of Electrical and Electronics Engineers, Inc.; 1984.

3. Beisel WR, Talbot JM, editors. *Research opportunities on immunocompetence in space.* Bethesda: Federation of American Societies For Experimental Biology; 1885.

4. Betts RF, Douglas JR. Influenza Virus. In: *Infectious Diseases, 3rd Edition.* Mandell GL, Douglas, Jr RG, Bennett JE, editors. New York: Churchill Livingstone; 1990, pp 1306-1325.

5. Blomberg DJ, Guth JL, Fattu JM, Patrick EA. Evaluation of a new classification system for anemias using consult learning system. *Comput Methods Programs Biomed* 1986; 22:119-125.

6. Bratko I, Mozetic I, Lavrac N. *KARDIO: A Study in Deep and Qualitative Knowledge for Expert Systems.* Massachusetts: The MIT Press; 1989.

7. Buchanan BG, Shortliffe EH. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Menlo Park, California: Addison-Wesley Publishing Company; 1984.

8. Clover RD, Abell T, Becker LA, Crawford S, Ramsey CN. Family functioning and stress as predictors of influenza B infection. *J Fam Pract* 1989; 28(5):535-539.

9. Culbert C. *CLIPS Reference Manual.* Houston: National Aeronautics and Space Administration; 1989.

10. Fry RJM. Radiobiological features of the space radiation environment. In: *Guidance On Radiation Received In Space Activities*. Bethesda, MD: National Council on radiation Protection and Measurements; 1989, pp 50-144.

11. Geissman JR, Schultz RD. Verification & validation of expert systems. In: *Validating And Verifying Knowledge-Based Systems*. Gupta U, editor. Washington: IEEE Computer Society Press; 1991, pp 12-25.

12. Giarrantano J. *CLIPS User's Guide*. Houston: National Aeronautics and Space Administration; 1989.

13. Giarrantano J, Riley G. *Expert Systems*. Boston: PWS-KENT; 1989.

14. Glezen WP, Decker M, Joseph SW, Mercready JR. Acute respiratory disease associated with influenza epidemics in Houston, 1981-1983. *J Infect Dis* 1987; 155(6):1119-1126.

15. Glezen WP, Grose N, Haddock A, Couch RB. Chemotherapy and management of respiratory virus infections. *Proceedings of the 16th International Congress of Chemotherapy*. 1989, pp 659.1-659.3.

16. Gmunder FK, Lorenzi G, Bechler B, Joller P, Muller J, Ziegler WH, Cogoli A. Effect of long-term physical exercise on lymphocyte reactivity: similarity to spaceflight reactions. *Aviat Space Environ Med* 1988; 59:146-151.

17. Goldberg J, et al. *A Strategy for Space Biology*. Washington, D.C.: National Academic Press; 1987.

18. Gupta U. *Validating and Verifying Knowledge-Based Systems*. Washington: IEEE Computer Society Press; 1991.

19. Haley RW, Aber RC, Bennett JW. Surveillance of nosocomial infections. In: *Hospital Infections, 2nd Edition*. Bennett JV, Brachman PS, editors. Boston: Little, Brown & Co.; 1986, pp 51-71.

20. Hand DJ. Artificial intelligence and medicine. *J Royal Soc Med* 1987; 80:563-565.

21. Kanas N. Psychosocial factors affecting simulated and actual space missions. *Aviat Space Environ Med* 1985; 56:806-811.

22. Keast D, Cameron K, Morton AR. Exercise and the Immune Response. *Sports Med* 1988; 5:248-267.

23. Layde PM, Engelberg AL, Dobbs HI, Curtis AC, Craven RB, Graitcer PL, Sedmak GV, Erickson JD, Noble GR. Outbreak of influenza A/ USSR/77 at Marquette University. *J Infect Dis* 1980; 142(3): 347-352.

24. Longini Jr. IM, Koopman JS, Haber M, Cotsonis GA. Statistical inference for infectious diseases. Risk-specific household and community transmission parameters. *Am J Epidemiol* 1988; 128(4): 845-859.

25. Marcot B. Testing your knowledge base. In: *Validating And Verifying Knowledge-Based Systems*. Gupta U, editor. Washington: IEEE Computer Society Press; 1991, pp 188-199.

26. Miller PL, Fisher PR. Causal models for medical artificial intelligence. In: *Selected Topics in Medical Artificial Intelligence*. Miller PL, editor. New York: Springer-Verlag; 1988, pp 11-24.

27. Moser MR, Bender TR, Margolis HS, Noble GR, Kendal AP, Ritter DG. An outbreak of influenza aboard a commercial airliner. *J Epidemiol* 1979; 110(1):1-6.

28. Nicogossian AE, Garshnek V. Historical perspectives. Nicogossian AE, editor. In: *Space Physiology and Medicine*. Philadelphia: Lea and Febiger; 1989, pp 17-29.

29. Nicogossian AE, Huntoon CL, Pool SL. *Space Physiology and Medicine*. Philadelphia: Lea and Febiger; 1989.

30. Nieman DC, Johanssen LM, Lee JW. Infectious episodes in runners before and after a roadrace. *J Sports Med Phys Fitness* 1989; 29(3):289-296.

31. Nieman DC, Nehlsen-Cannarella SL. The immune system. In: *In: Principles and Practice of Sports Medicine*. Rians CB, editor. Champaign, IL: Human Kinetics Publishers; 1991, pp 1-49.

32. Olson JG, Ksiazek TG, Irving GS, Rendin RW. An explosive outbreak of influenza caused by a/USSR/77-like virus on a United States naval ship. *Milit Med* 1979; :743-745.

33. Pierson DL. *Space Station Infectious Disease Risks*. Houston: National Aeronautics and Space Administration; 1986:1-1 to 4-16.

34. Pons VG, Canter J, Dolin R. Influenza A/USSR/77 (H1N1) on a university campus. *Am J Eiodemiol* 1980; 111(1):23-30.

35. Reuman PD, Bernstein DI, Keely SP, Sherwood JR, Young EC, Schiff GM. Influenza-specific ELISA IgA and IgG predict severity of influenza disease in subjects prescreened with hemaglutination inhibition. *Antiviral Res* 1990; 13:103-110.

36. Riley G. *CLIPS Architecture Manual*. Houston: National Aeronautics and Space Administration; 1989.

37. Robbins F, et al. *Exploring the Living Universe/A Strategy for Space Life Sciences*. Washington, D.C.: National Aeronautics and Space Administration; 1988.

38. Taylor R, Nemaia H, Tukuitonga C, Kennett M, White J, Rodger S, Levy S, Gust I. An epidemic of influenza in the population of Niue. *J Med Virol* 1985; 16:127-136.

39. *Vital and Health Statistics: Current Estimates From the National* Health Interview Survey, 1988. Hyattsville, Maryland: DHHS Publication; 1989.

40. Winkel P. The application of expert systems in the clinical laboratory. *Clin Chem* 1989; 35(8):1595-1600.

41. Woodson WE. *Human Factors Design Handbook*. New York: McGraw-Hill Book Co.; 1981.

Session H4: HUMAN MACHINE INTERACTIONS

Session Chair: Dr. Arthur Beller

DESIGN FOR INTERACTION BETWEEN HUMANS AND INTELLIGENT SYSTEMS DURING REAL-TIME FAULT MANAGEMENT

Jane T. Malin
Johnson Space Center
Houston, TX 77058
malin @ aio. jsc. nasa. gov.

Debra L. Schreckenghost
TheMITRECorporation
1120 NASA Road One
Houston, TX 77058
schreck @ aio. jsc. nasa. gov.

Carroll G.Thronesbery
The MITRE Corporation
1120 NASA Road One
Houston, TX 77058
tbery @ aio. jsc. nasa. gov.

ABSTRACT

Initial results are reported from interdisciplinary projects to provide guidance and assistance for designers of intelligent systems and their human interfaces. The objective is to achieve more effective human-computer interaction (HCI) for real-time fault management support systems. Studies of the development of intelligent fault management systems within NASA have resulted in a new perspective of the user. If the user is viewed as one of the subsystems in a heterogeneous, distributed system, system design becomes the design of a flexible architecture for accomplishing system tasks with both human and computer agents. HCI requirements and design should be distinguished from user interface (displays and controls) requirements and design. Effective HCI design for multi-agent systems requires explicit identification of activities and information that support coordination and communication between agents. We characterize the effects of HCI design on overall system design and identify approaches to addressing HCI requirements in system design. Our results include definition of (1) guidance based on information-level requirements analysis of HCI, (2) high-level requirements for a design methodology that integrates the HCI perspective into system design, and (3) requirements for embedding HCI design tools into intelligent system development environments.

INTRODUCTION

Two multi-year, interdisciplinary projects are currently in progress to develop technology that helps developers of intelligent systems with real-time fault management capabilities achieve more effective human-computer interaction (HCI) design. The objective of the work is to specify requirements for methods and tools and to provide design guidance to assist developers in making intelligent systems better team players. Intelligent systems can be viewed as computer agents that share the task of process monitoring and control with human agents. An important part of designing the intelligent system is to specify how these human and computer agents will interact and what information they will exchange to be an effective fault management team.

Mitchell (1987) has observed that functional requirements for complex systems rarely specify information needed to support tasks of human controllers or operators. Yet, real-time interactive process control software usually requires that the needed information be easily accessible and discernible. This is especially true when this software includes intelligent systems to provide advanced support for human operators. Intelligent systems can provide additional information at a higher level of abstraction (e.g., reasoning about faults in addition to fault signatures in numeric data), but can also increase the need to coordinate and to monitor the software. Lack of integration of HCI information requirements into the functional requirements specification is an important problem in the design of real-time interactive process control software. It is even more critical when such software includes an intelligent system component.

However, user interface (UI) and HCI analysis and design has traditionally been viewed as relevant not to the functional requirements, but rather to the non-functional constraints (Roman, 1985), which are often applied later in software development. As Marshall (1991) phrases it, there is a common perception that human factors specialists "should be brought in to sprinkle magic dust on the interface or workstation once it is

largely developed." But deferring analysis of all types of HCI needs until late in system development contributes to lack of system support for the operator tasks. HCI considerations, however, can affect what the intelligent system does (i.e., its functionality) as well as constrain how the intelligent system is implemented.

The problem is due partly to a confusion between HCI requirements analysis and UI design. Human-computer interaction requirements refer to the information exchange between the user and the computer when performing the task which the software system is intended to support. The user interface design, on the other hand, concerns the display and control software and hardware that are the media for information presentation and dialog with the user. One of the dangers of confusing HCI and UI is the tendency to overlook task-level information needs (e.g., information needed to make a good decision), while focusing on physical interface characteristics (e.g., icon shapes and colors) and low-level interface functions (e.g., radio buttons and pull-down menus). This danger is described in detail by Woods and Eastman (1989) while introducing their "levels" of display design approach.

Specification of HCI requirements is critical in design of real-time, interactive intelligent systems. Focusing on screen display and low-level interaction sequences can lead to ignoring the more global issues of whether the right information is flowing between the software system and the user at the right time. Therefore, it is important to clarify the UI/HCI distinction and tackle the problem of including consideration of HCI information exchange requirements in development methods and tools.

METHODS

In one of the two projects, intelligent fault management systems within NASA have been evaluated for insight into the design of systems with complex HCI (Malin et al., 1991). Fifteen intelligent systems were selected from a variety of aerospace programs, including Space Shuttle, Space Station, unmanned spacecraft, and military and advanced aircraft. Information was collected by interviewing intelligent system developers and users, observing intelligent system demonstrations, and referencing available documentation about applications. In one case, participation in the design of the Space Shuttle Payload Deployment and Retrieval System (PDRS) Decision Support System (DESSY)

project permitted first-hand observation of the design process and provided an opportunity to test design recommendations and develop examples of design guidance.

The other project has focused on conceptual design and prototyping of methods and tools for development of intelligent systems and their user interfaces. A prototype software toolkit, User-Intelligent System Interface Toolkit (UISIT), has been developed to support a methodology for developing conceptual designs for interactive intelligent systems and their user interfaces. An evaluation of UISIT prototype tools and methods is now being conducted in an application case. UISIT-based methods and tools are being used to support the PDRS DESSY project, to develop a prototype intelligent fault management support system. The purpose of this effort is to evaluate the underlying HCI development methodology and its support by the tools.

RESULTS

Preliminary results of this work indicate that the concept of information requirements is the productive focus for identifying needed modifications and additions to traditional methods, guidelines and tools. Information requirements identify the information which must be exchanged between the user and the intelligent system in order to support user tasks. Specifically addressing information exchange requirements helps the designer to elaborate what capabilities the intelligent system and the UI will need to have before determining exactly how to implement these capabilities. As such, these are an important subset of the external interface requirements specification. Information requirements also help to coordinate the design of the user interface and the application software.

Our solutions to the problems outlined above can be partitioned into the following areas, which outline what it will take elevate HCI requirements to the level of functional requirements: 1) HCI Design Guidance for Information Requirements, 2) Development Methodology for Information Requirements, 3) User Interface Tools and Methods for Information Requirements.

1. HCI Design Guidance for Information Requirements

Current forms of guidance should be extended to support the development of information exchange requirements.

Traditional guidance focuses primarily on visual appearance and style, offering little assistance in designing an intelligent system and its user interface which provides the right information to support user task performance. The example in Figures 1 and 2 illustrates this new guidance and its emphasis on information exchanged between the human and the computer. This guidance assists designers in determining the information which is required from the intelligent system for the display. It is relevant to decisions made early in the development process that constrain the information available for display. Unlike decisions about how to display information, they can not be deferred until later in the development process. This design guidance can help the HCI designer be more effective in identifying requirements for both the intelligent system and the user interface. Early involvement of the HCI designer in intelligent system development helps to integrate the efforts of the HCI designer and the intelligent system designer.

2. Development Methodology for Information Requirements

Another approach to elevating the consideration of HCI information exchange requirements is to provide an intelligent system development methodology that incorporates HCI considerations as an integral part of design. Such a methodology features the explicit specification of information requirements, including application of the new guidance identified above.

Initial research efforts (Johns, 1990; Malin et al., 1991) have resulted in a new perspective of the user -- the user as another type of agent in a heterogeneous, cooperating, distributed system. System design then becomes the design of an architecture for accomplishing domain tasks with the available human (i.e., users) and computer (i.e., applications) agents. HCI considerations are an important part of such system design, even before user interface design is addressed. We have defined the following stages in developing such a system (illustrated in Figure 3):

- *Description of domain tasks.* Monitoring and fault management tasks, whether performed by humans or software, are described in terms of goals and the actions required to achieve these goals. These goals and actions need not yet be assigned to specific types of agents.

- *Identification of resources provided to perform tasks and the constraints that affect task performance.* Resources include the capabilities of the operational environment and of the agents, and the availability of information. Constraints result from complexity, dynamics, and deficiencies in these resources.

- *Specification of agent activities and valid agent behaviors in an architecture for multi-tasking and dynamic task allocation.* Actions defined in the task description are assigned as specific agent activities consistent with the available resources and the inherent constraints of the fault management team and the operational environment. The task description is not complete until it includes activities supporting both domain task performance and coordination between agents of the fault management team. Note the implications for modification of task analysis and task description techniques to capture human-computer coordination activities.

- *Evaluation of the activity specification using expected operational scenarios.* The specification of agent activities is evaluated using complex, realistic activity sequences and modes of agent interaction to derive the information and functionality requirements. These scenarios should include tasks shared by multiple agents supporting both agent coordination and control of the monitored process, they should include both nominal and anomalous activity sequences, and they should highlight information used and exchanged to support each task.

- *Analysis of information exchange requirements.* Requirements for information are explicitly identified for both the intelligent system and the user interface, based on results of the evaluation using operational scenarios.

3. User Interface Tools and Methods for Information Requirements

A third way to promote the consideration of HCI information exchange requirements involves user interface tools and methods. User interface tools should support:
- Information requirements development
- Design team communication
- Run-time software module communication

User interface tools should support information requirements development and use as well as user interface design and management. Designing the HCI so that the end user always has the right

532

EXAMPLE OF HCI DESIGN GUIDANCE

Topic: Intervention Into Intelligent System Reasoning Process

Problem:

An intelligent system can become "disoriented" and require the operator to redirect it onto a more productive
path of investigation (e.g., the system investigating an unproductive hypothesis or failing to investigate a likely hypothesis).

Recommendation:

Operator intervention into the reasoning process of the intelligent system can be used to manage errors in the intelligent system.
Methods of intervention include modification of the reasoning process or selection of an alternate reasoning mechanism.

Supporting Information:

Intervention into intelligent system processing is one approach for redirecting a "disoriented" intelligent system onto a more productive
path of investigation. The intelligent system must be designed, however, to permit such intervention. To effectively intervene in
processing, the operator must first have a good understanding of the reasoning strategy used by the intelligent system.

Example Illustrating Problem:

See the following page for an example that illustrates the problem.

Example Illustrating Recommendation:

See the following page for an example that illustrates the recommendation.

Techniques:

Methods of intervening in the method of processing information include modification of the reasoning process or selection of an alternate
reasoning mechanism. Examples of modification of the reasoning process are (1) setting processing priorities (e.g., what hypothesis
to investigate first), (2) alteration of hypotheses, and (3) specification of alternate solutions.

Cross References:

The user interface used to intervene should reinforce the operator's understanding of the reasoning process to assist in identifying
appropriate action. See section 4.1.2 for a discussion of presenting information about intelligent system reasoning. See the discussion
of risk in intervening in the intelligent system later in this section.

Research Issues:

Making intelligent systems easily interruptable and redirectable, developing vocabularies for communicating advice about control
decisions which reduce the amount of knowledge required by human team members about intelligent system internals.

Figure 1. Example of HCI Design Guidance, Page 1.

EXAMPLE OF HCI GUIDANCE

Topic: Intervention Into Intelligent System Reasoning

AGENT INTERACTION ILLUSTRATING PROBLEM

In this case, the operator is unable to communicate
information about a likely fault and its correction
procedure to the intelligent system. Unaware of this
information, the intelligent system would continue to
pursue fault isolation on an incorrect set of suspected
faults, even when the actual fault was corrected. The
operator would have to do a restart to reset the knowledge
base and "fix" the intelligent system. Figure illustrates
the agent interaction.

Illustration: Operator Restarting Intelligent System

AGENT INTERACTION IIIUSTRATING RECOMMENDATION

In this case, the ability to affect reasoning by altering
hypotheses used by the intelligent system has been provided
to the operator. The operator "informs" the intelligent
system about the fault, performs a corrective procedure,
and normal processing to continues. Figure illustrates the
agent interaction.
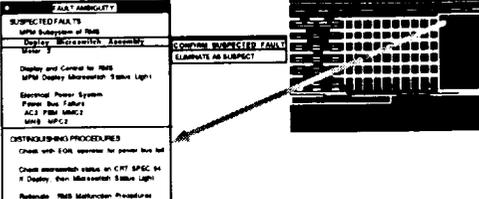
Illustration: Operator Informing Intelligent System of a Fault

Figure 2. Example of HCI Design Guidance, Page 2.
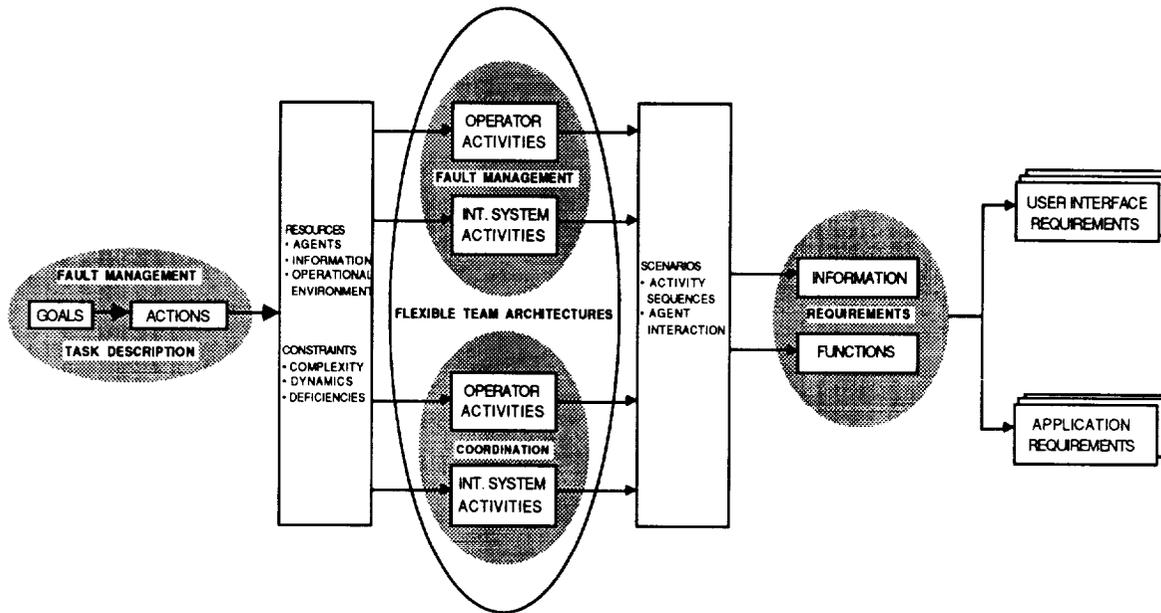
Figure 3. Recommended Requirements Development Methodology.

information at the right time during task performance is difficult, and it requires different tools from those which support user interface design. Tools should support the representation of information exchange items and the interrelationships among these items.

User interface tools should also use information exchange requirements for coordination and communication among design team members. This allows early and continued consideration of the information exchanged between the user interface and the applications components.

The preliminary design methodology developed in conjunction with the UISIT prototype toolkit is illustrated in Figure 4. It defines information requirements early in the system design process and supports coordination among design team members. Analysis and preliminary design occur first to derive high-level descriptions of intelligent system functions, tasks, architecture, and HCI. From these descriptions, requirements are derived for the major software component functions and for HCI information exchange. Using information requirements, the design team can be partitioned into smaller groups, with each

group performing a separate but coordinated and concurrent design of a software component. As it is discovered that requirements must change to match constraints, the information requirements can serve as a point of coordination among the design team members. For instance, if the intelligent system designer discovers that additional information from the user is required to complete an activity, the information exchange specification serves as a single point of coordination with the user interface designer. The intelligent system designer and user interface designer can prototype their software modules separately while using the information exchange specification to coordinate with one another.

User interface tools and methods should support user interface prototyping. An especially effective form of prototyping support is to represent information requirements in the same module which supports run-time information exchange. This allows the separate prototyping of the user interface and the applications modules. It also requires the information exchange specification to be stated unambiguously.

534
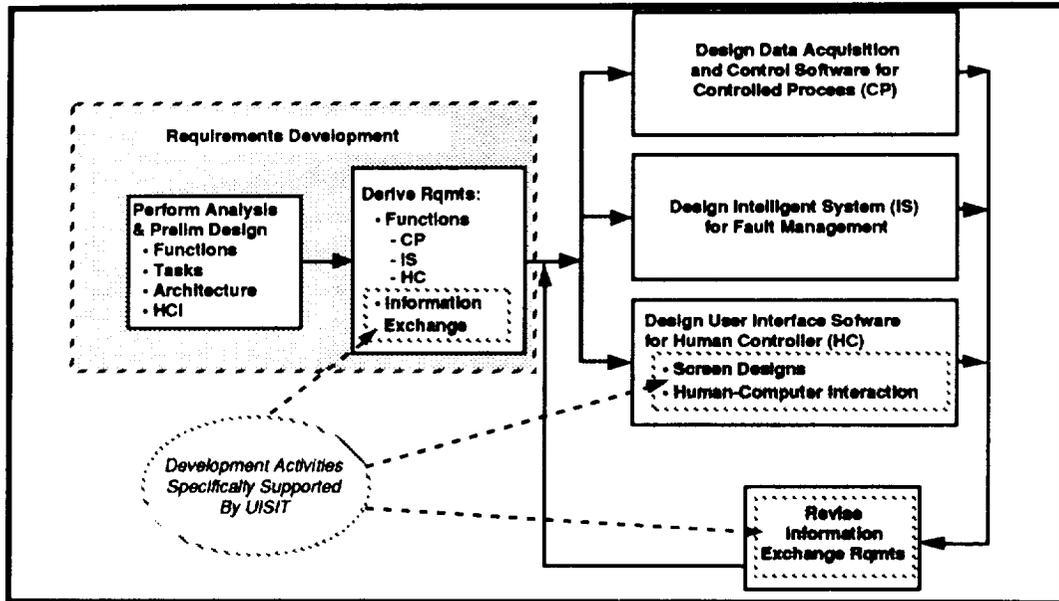
Figure 4. Recommended UISIT Methodology.

Finally, it presents a design team member with an external cue to coordinate changes with the remainder of the design team.

Figure 5 illustrates how a software system built with the aid of UISIT segregates the information exchange specification from the specification of the intelligent system and the data acquisition software. This segregation allows each specialist to apply unique expertise to a single software component, while providing a means for coordinating design team efforts through the information specification module. When it is discovered that a change needs to be made to the information specification module, it is clear to all involved that the change needs to be coordinated with the other members of the development team. In this way, the information specification module serves as an abstraction of what each module designer needs to know about the design of the other modules.

Figure 6 shows an object hierarchy comprising the information specification component of a sample system built for UISIT evaluation.

The contents of the object hierarchy represent the information requirements and their relationships for PDRS DESSY. The object layer was constructed within the UISIT framework, in accordance to the suggested UISIT hierarchy. Because it is encoded as an object layer, supporting run-time communications between the user interface software and the applications software, the information specification module is unlikely to be misinterpreted by members of the development team

RELATED WORK

In 1983, Norman outlined four strategies for improving to human-computer interaction design. Our recommendations for enhancing these strategies are summarized below:

- *Help the designer to maintain an awareness of the user's needs.* Rather than simply trying "to impress upon the designer the seriousness of the matter," we propose that it would be more effective to assist in identifying user needs by elevating analysis of information exchange

535
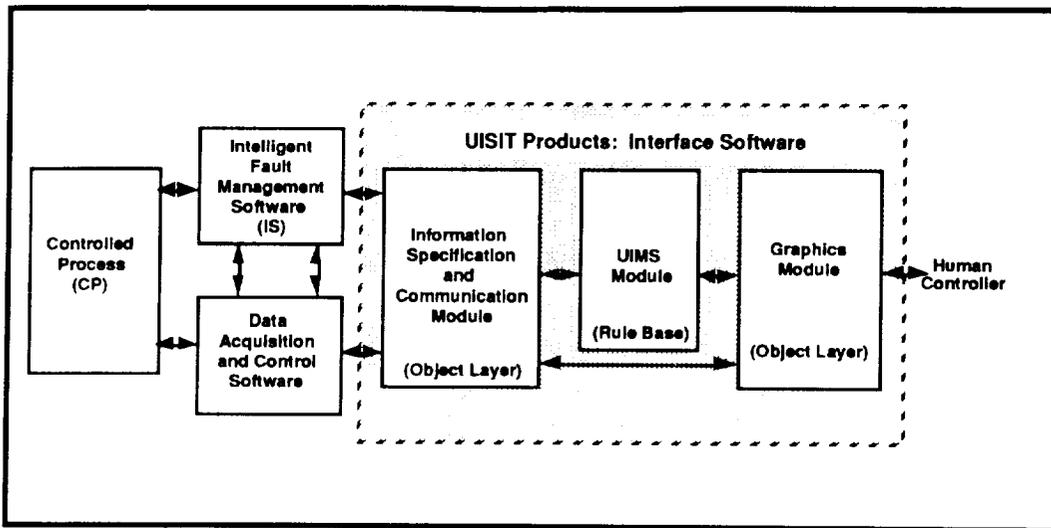
Figure 5. UISIT-Supported Architecture.

between the software and the user to a functional requirements level.

• *Provide the designer with methods and guidelines.* Most assistance has concentrated on support for evaluation of user interface designs. We shift the focus to support for analysis of information needed to support user tasks.

• *Provide software tools for interface design.* As Norman suggests, this is a natural follow-on to providing methods and guidelines. The software tools can provide guidance in two ways: (1) explicitly in the form of on-line guidelines, templates, and examples; and (2) implicitly by the design of the toolkit, which requires specific architectural components and a specific sequence of developmental steps. In either case, such tools make it easier for the designer to follow the prescribed methods and guidelines.

• *Separate the interface design from other programming tasks.* While Norman's original suggestion was based on such principles as software reusability and interface modifiability,

we have found that this practice can also support design team communication.

While we have been emphasizing HCI development in our own work, we do not intend to imply that user interface issues are unimportant. We emphasize HCI development because it is too often neglected. Both HCI and UI challenges are important, and both are necessary for effective system design. Without effective HCI design, the UI designer may not have a clear idea of the information exchange goals he is trying to optimize, and the resulting software may not support the user's task performance. Without effective UI design, the user interface may not convey the information effectively, thereby detracting from the support of the user's task performance. Since both types of issues are necessary for effective design and since HCI issues are often overlooked, we have concentrated our work on HCI issues. There are significant UI design challenges in monitor and control of complex processes, due to the large quantities of real-time information, accompanied by deficiencies of the quality and availability of this information.

536

Figure 6. Object Hierarchy and Sample Definition for UISIT Communication Layer.

Because the conclusions of this report are based on the study of intelligent systems for real-time fault management, the reader may be concerned about how well they generalize to other types of software systems, especially in view of Leveson's (1990) warning that applying common data processing approaches to process control systems can lead to disaster. We believe that our results for HCI design guidance are applicable to *complex* software systems. Complex software systems are characterized by (1) large amounts of information from multiple sources, (2) sophisticated software capability, often with multiple tasks performed concurrently or jointly by human and computer, (3) time-constrained processing with deficiencies in information quality and availability, and (4) active information exchange between human and computer. Our results on development methodology and tools appear to be generalizable to a broader set of systems, in particular systems where information exchange between human and computer represents a significant aspect of using the software.

SUMMARY

This paper has introduced an on-going research project to improve human-computer interaction for fault management intelligent systems. This issue is being investigated by means of a case study, by participation in intelligent system design, and by prototyping methods and tools. Preliminary results indicate that elevating HCI information exchange requirements to the level of software functional requirements is critical to designing software which supports user task performance. The following is a summary of the findings from this research effort:

1.   HCI Design Guidance for Information Requirements
•   User interface design guidance, which focuses on visual appearance and style of information presentation, should be extended HCI, to also assist designers in developing information requirements.
•   HCI design guidance should be integrated within a development methodology that supports use of these guidelines.

537

2. Development Methodology for Information Requirements
• Task analysis and task description techniques should be modified to identify human-intelligent system coordination activities and to support identification of information requirements.
• Developers should adopt a methodology that makes guidance easier to use and integral to the development process from the early stages of analysis and design.
• Mechanisms for communication of information requirements and coordination of design activities among members of the development team should be provided with the methodology.

3. User Interface Tools and Methods for Information Requirements
• User interface tools should support information requirements development and should include explicit representation of information items and the interrelationships among them.
• User interface tools should provide information requirements as a point of coordination and communication among members of development team.
• User interface tools should represent information requirements in the same module that supports run-time information exchange.

REFERENCES

Johns, G. (1990, September). *Graphic Interfaces to Intelligent Fault Management Systems: Issues and Guidelines* (Report No. MTR-90W00103). McLean, VA: The MITRE Corporation.

Leveson, N. G. (1990, November). "The Challenge of Building Process-Control Software." *IEEE Software, 7*(6), 55-62.

Marshall, C. (1991, March). "Ergonomics Is Dead: Long Live Ergonomics." *Human Factors Society Bulletin, 34* (3), 4-6.

Malin, J. T., Schreckenghost, D. L., Woods, D. D., Potter, S. S., Johannesen, L., Holloway, M., & Forbus, K. D. (1991, July). *Making Intelligent Systems Team Players: Case Studies and Design Issues, Volume 1. Human-Computer Interaction Design* (NASA Technical Memorandum). Houston, TX: NASA-Johnson Space Center.

Mitchell, C. M. (1987). "GT-MSOCC: A Domain for Research on Human-Computer Interaction and Decision Aiding in Supervisory Control Systems." *IEEE Transactions on Systems, Man, and Cybernetics, 17*(4), 553-572.

Norman, D. A. (1983, December). "Design Principles for Human-Computer Interfaces." *CHI '83 Proceedings*. New York: ACM, 1-10.

Roman, G. C. (1985, April). "A Taxonomy of Current Issues in Requirements Engineering," *IEEE Computer, 18*(4), 14-21.

Woods, D. D., & Eastman, M. C. (1989). "Integrating Principles for Human-Computer Interaction into the Design Process: Heterarchically Organized HCI Principles." In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE.

I

A HUMAN FACTORS EVALUATION OF THE ROBOTIC INTERFACE FOR SPACE STATION *FREEDOM* ORBITAL REPLACEABLE UNITS

Carlos E. Sampaio[1], Ellen Y. Hwang[1], Terence F. Fleming[1],
Mark A. Stuart[1], and A. Jay Legendre[2]
[1]Lockheed Engineering and Sciences Company
Houston, TX 77058
[2]NASA Johnson Space Center
Houston, TX 77058

ABSTRACT

An orbital replaceable unit (ORU) is often defined as any orbital unit aboard Space Station with a wearout life of less than 30 years. The capability of successful changeout of these units by remote manipulation is critical to the ORU to telerobot interface design. A human factors evaluation of the selected interface showed certain inadequacies of the alignment target concept that was part of the interface package. Alternative target concepts which addressed these inadequacies were developed and are presented in this paper. Recommendations from this work will be incorporated into NASA requirements documents which ORU suppliers and manufactures must then build to.

ACRONYM LIST

IDR	Interface Design Review
JSC	Johnson Space Center
ORU	Orbital Replaceable Unit
ROIL	Remote Operator Interaction Laboratory
RSIS	Robotic Systems Integration Standards
SSF	Space Station *Freedom*

INTRODUCTION

The changeout of orbital replaceable units (ORU) will be a vital day to day activity for the crew aboard Space Station *Freedom* (SSF). Performing the changeout of these units by both the extravehicular astronaut as well by the various robotic devices aboard SSF will be critical to the success of *Freedom's* operation. Therefore it is critical that the interface to these ORUs assure successful manipulation. In response to this, an Interface Design Review (IDR) process has been formed in order to assess SSF robotic interfaces and to establish appropriate interface design standards which will be published in the Robotic Systems Integration Standards (RSIS). These will become the standards which ORU suppliers and manufacturers must build to.

As a result of the IDR process, the robot to ORU interface package selected for ORUs of 1200 pounds or less was the Spar Aerospace Limited design. Once the selection was made, a design validation process began with the intent of evaluating the suitability of the selected interface design. This entailed making recommendations intended to refine and improve the Spar package. Several NASA as well as private laboratories were chosen to participate in the design validation process, among them the Remote Operator Interaction Laboratory (ROIL) at NASA's Johnson Space Center (JSC) in Houston. The ROIL was chosen specifically for its expertise in dealing with human factors issues relevant to the interface design.

In evaluating the Spar package, personnel at the ROIL determined that the Spar ORU alignment target had the greatest impact from a human factors perspective. Consequently, it was this part of the interface which ROIL personnel chose to focus on in an attempt to refine it. This paper discusses the approach taken by the ROIL during its evaluation and refinement of the Spar alignment target.

THE SPAR ALIGNMENT TARGET

Shown graphically in Figure 1, the Spar alignment target was 2 inches by 2 inches square. The front and base plates of the target were separated by a depth of .5 inch. This .5 inch separation between the .25 inch diameter white circle on the face plate and the .284 inch diameter black circle on the base plate provided the cues necessary to align the pitch and yaw rotational axes. This was done by making sure a black ring from the base circle was always visible around the white circle on the face plate. An overlay displayed on the end effector camera view, when aligned properly with the target, provided the cues necessary to adjust the additional rotational roll axis as well as the translational X, Y, and Z axes. Once aligned properly, the target—mounted 2.5 inches above the center of the grapple fixture on the ORU—was designed to assure that the misalignment tolerances of the grapple fixture were met before a grapple attempt was made.
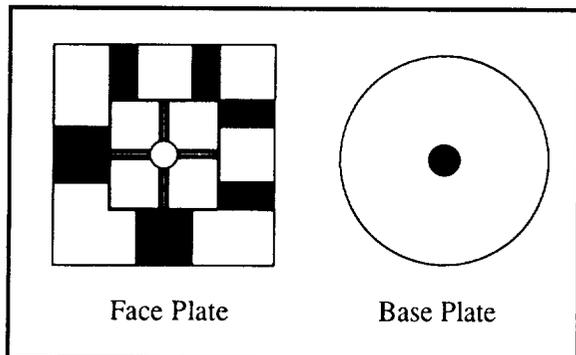

Figure 1. Face plate and base plate concepts incorporated into the original Spar target.

Certain aspects of the Spar alignment target design were very well conceived. The compactness of the target allowed for it to be mounted appropriately near each grapple point. Further, because it was an enclosed unit, it would pose little to no risk of snagging an extravehicular astronaut. The simplicity of the target was also a good feature. Corrective alignments were thus logical and easy to interpret. Initial evaluations by the ROIL, however, pointed out several shortcomings inherent to the target's design.

First among these limitations was a matter of geometry. According to the operational scenario provided by Spar, at grapple the camera to target distance was defined to be 4 inches. Figure 2 shows that according to the dimensions specified by Spar, the minimum distance at which the black circle on the base plate would be visible around the white circle on the face plate was 3.67 inches. As a result, at the 4 inch grapple distance, the cues provided by the base circle were so slight that final alignment corrections were often difficult to determine.

Formula 1

$$\frac{R1}{D} = \frac{R2}{D + 0.5"}$$

Spar Dimensions
R1 = 0.125"
R2 = 0.142"
D = 3.67"
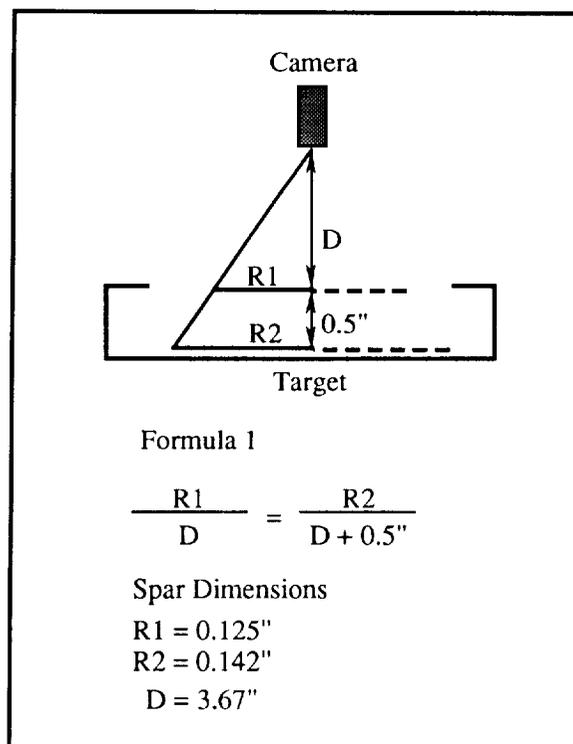
Figure 2. The relevant target and camera dimensions (not to scale) are represented in Formula 1. Given the specified Spar dimensions, the minimum distance necessary to view the base circle around the face plate circle was 3.67 inches.

Accuracy evaluations were performed by mounting the target to a rotating base 4 inches in front of and centered with the camera. An operator would then call out when they felt that a zero degree rotational alignment (i.e. perfect alignment) had been achieved. Due to the geometry of the Spar target's relevant features,

540

accuracy was typically within 2° to 3° to either side of zero degree rotational alignment.

When examined under controlled lighting conditions, certain other drawbacks to the Spar alignment target design became apparent. When a collimated light source—in this case a solar simulator—was directed at the target from an angle displaced 15° to 25° from the camera centerline, certain angular displacements of the target caused the face plate to obscure the black base plate circle by shadows from either the outside frame of the face plate or by the white circle on the face plate. The corrective action needed in these cases would not be clear and could result in the issuance of an incorrect command.

Similarly, at certain incident angles, light would reflect off the highly specular paint specified by Spar to the point that features of the target were completely washed out due to the blooming effect that would appear on the video monitor. If the iris on the camera's lens were closed to the point that the blooming effect was relieved, the contrast would be so slight that the target's features were still indistinguishable on the monitor view.

PROPOSED TARGET MODIFICATIONS

It was clear that the desirable features of the Spar target concept needed to be preserved in any new concept to be proposed. Just as clear, however, was the need to address the limitations of the original Spar target's design. With this in mind, ROIL personnel began an iterative process intended to devise target concepts which preserved the desirable features of the Spar design, but which also improved upon its limitations. It should be noted that none of these proposed target modifications addressed the reflectivity of the paint due to the long turnaround time in having a target concept painted to specification. Paint reflectivity will be addressed as an issue separate from features of the target design.

Rather than address each limitation of the Spar design separately, alternative designs were developed in order to address the various issues as a system rather than as independent problems. This way, solutions which addressed one potential flaw in the Spar target design would be less likely to complicate another. This was done specifically with respect to the shadow and geometric problems encountered in the Spar design.

Figure 3 depicts an alternative target design, designated concept A, which both reduced the likelihood of any potential confusion caused by shadows as well as assuring that all alignment cues are clearly discernable at any camera to target distance. The opening in the face plate was widened to both allow more light into the target as well as reduce the likelihood of a shadow from the edge of the face plate obscuring the center of the base plate. Likewise, the width of the crosshairs were reduced from the original Spar dimension of .12 inch to .02 inch. The narrower crosshair created much less of a shadow on the target base, eliminating the possibility of obscuring relevant features on the base plate. Further, the alternating black and white colors on the crosshairs, in combination with the features of the base plate, made them always distinguishable from the background. This made corrective movements easier to determine. Consequently, the same accuracy evaluations as performed with the original Spar target resulted in accuracy typically within .33° to either side of zero degree rotational alignment with no apparent shadowing problems.
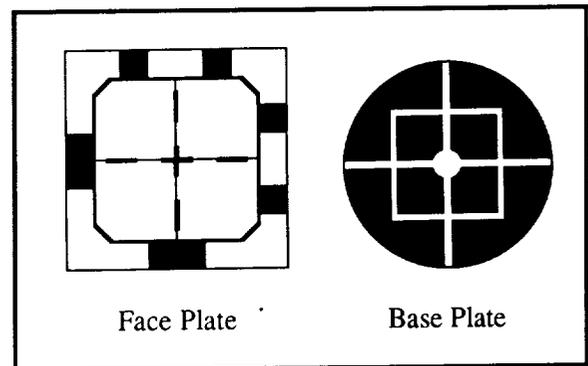
**Figure 3.** Face plate and base plate depictions proposed in alternative target concept A.

Certain concerns were raised regarding target concept A, however. Chief among them was a lack of robustness with respect to the crosshairs. Any contact with the crosshairs by a robotic end effector or by an extravehicular astronaut could potentially bend or even break them. Therefore, further designs were devel-

541

oped with the intent of addressing that concern in particular.

It was felt that the same level of accuracy could be maintained if the crosshairs were widened. Preserving many of the features of target concept A with widened crosshairs could result in a more robust target that was easy to interpret and which could provide the cues necessary to achieve the accuracy of the first alternative design. At the same time, it was felt that additional cues could be built into the target which would provide cues redundant to the primary alignment cue, in this case the center of the crosshair.

Figure 4 depicts target concept B. The crosshairs have been widened to .063 inch, making them very robust. At the same time, the level of accuracy exhibited by concept A was preserved in this concept. Once again, accuracy evaluations resulted in alignment typically within .33° to either side of zero degree rotational alignment and again, no apparent shadowing problems appeared. The new feature incorporated into this target was a redundancy of the alignment cues used prior to grapple. To achieve proper alignment of the end effector at the 4 inch camera to target distance, the center cross on the face plate crosshairs had to be centered within the white circle on the base plate. The redundant cues were provided by the black hashmarks further out on the crosshairs. When proper alignment was achieved at the 4 inch distance, the inside edge of the black hashmarks appeared to touch the outside edge of the base plate black circle. Pitch and yaw cues which the target must provide are then also offered by these redundant cues as well.

The redundant cues offered by this concept do two things. First, they provide an extra measure of certainty that proper alignment has been achieved prior to attempting grapple. If both sets of alignment cues tell the operator that the end effector is positioned properly, the operator will be that much more sure at the time of grapple. The second, and perhaps more important feature offered by the redundant cues is that they can used as primary alignment cues in the event that the original cues, i.e. the center of the target, cannot be used for alignment. This scenario could occur if certain harsh shadows extended over the tar


Face Plate   Base Plate
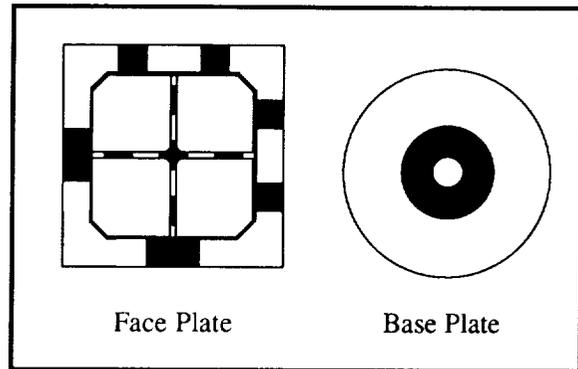
Figure 4. Face plate and base plate depictions proposed in alternative target concept B.

get, or if the target were partially damaged in some way. The redundant cues might allow a successful grapple to be made which might otherwise not be attempted due to the extraneous circumstances.

FUTURE WORK

The alternative designs presented here are only two of several iterations on the target concept conceived by ROIL personnel. While these represent the most promising of the iterations attempted so far, other concepts may offer even greater promise. Consequently, work on improving the design has not terminated and further iterations on the concept will proceed.

It is important to note that a clear line of communication has already been established between ROIL and Spar personnel. Thus, this work is becoming a collaborative effort consisting of feedback offered by the ROIL being strongly considered by personnel at Spar.

Further work is being planned by the ROIL to address the issues raised in this paper. Alternative paints will be evaluated at the ROIL with the intent of determining which have acceptably low reflectivity characteristics. The primary effort to be taken on by the ROIL, however, will be to perform an operational evaluation of these and other target concepts. Test subjects will consist of personnel at JSC with experience in operations, including the crew. Data will be gathered with respect to accuracy achieved during each alignment run as well as subjective data regarding how well each

542

subject felt they could interpret the targets evaluated. Results of these evaluations will be incorporated into the RSIS documents.

CONCLUSION

It seems clear that these proposed modifications are improvements upon the original Spar design. The line of communication opened between ROIL and Spar personnel as well as the incorporation of this work into RSIS documentation means that this work will have clear implications to future space hardware design. The end product of this work will hopefully result in an alignment target that is very easily interpretable and which will work under a wide variety of situations.

ACKNOWLEDGEMENTS

SITUATION AWARENESS IN COMMAND AND CONTROL SETTINGS

Robert W. Patterson, Ph.D.
Det 5, Armstrong Laboratory
Air Force Human Resources Laboratory
Wright-Patterson AFB, OH   45433

ABSTRACT

Many important military and civilian operational settings can be
described as command and control (C2) environments characterized by
high information load, extensive team coordination, and communica-
tion demands. To address these problems, designers have turned to
the application of automation and decision aids. Currently, there
are few techniques for assessing the performance of C2 operators
with which to guide applications. A need exists for constructs on
which to evaluate C2 systems. One possible concept for addressing
these issues is that of Situation Awareness (SA).

The SA concept and methodologies have developed largely around the
cockpit environment. SA describes how a pilot perceives and
comprehends the environment including aircraft states and aircraft
position. SA, then, may be used as a basis for gathering data to
evaluate operator and system performance.

This paper will discuss the extension of the SA concept to a team
context related to C2 environments. The paper will then discuss
the implications of such an approach for developing a methodology
to empirically measure team SA in a C2 environment which should
allow for the assessment of SA of both individual operators and
overall SA of operational teams. The discussion will also focus on
how a knowledge of SA may impact the design and development of
systems which optimize rather than maximize information access and
processing and, subsequently, improve C2 decision making.

EVALUATING HUMAN PERFORMANCE MODELING FOR SYSTEM ASSESSMENT: PROMISE AND PROBLEMS

Robert W. Patterson, Ph.D. and Capt Michael J. Young
Det 5, Armstrong Laboratory
Air Force Human Resources Laboratory
Wright-Patterson AFB, OH   45433

ABSTRACT

The Air Force Human Resources Laboratory is currently involved in the development and evaluation of computational human performance models. This development work is intended to develop models which can be used to interact with system prototypes and simulations to perform system assessment. Currently LR is working on a set of models emulating cognitive, psychomotor, auditory, and visual activity for multiple operator positions of a C2 simulation system. These models, developed in conjunction with BBN Systems and Technologies, function within the simulation environment and allow for both unmanned system assessment and manned (human-in-loop) assessment of system interface and team interactions. These are relatively generic models with built-in flexibility which allows modification of some model parameters.

These models have great potential for improving the efficiency and effectiveness of system design, test, and evaluation. However, the extent of the practical utility of these models is unclear. Initial verification efforts comparing model performance within the simulation to actual human operators on a similar, independent simulation have been performed and current efforts are directed at comparing human and model performance within the same simulation environment. The focus of this presentation will be on the on-going validation process through which the models will be tested and enhanced. Topics covered will include LR's approach; problems with model assessment both conceptual and methodological; and implications for this type of modeling approach.

Time Management Displays for Shuttle Countdown

The Intelligent Launch Decision Support System project at Kennedy Space Center
is developing a Time Management System (TMS) for the NASA Test Director (NTD)
to use for time management during Shuttle terminal countdown. TMS is being
developed in three phases: an information phase, a tool phase, and an advisor
phase. The information phase is an integrated display (TMID) of firing room
clocks, of graphic timelines with Ground Launch Sequencer events, and of
constraints. The tool phase is a what-if spreadsheet (TMWI) for devising
plans for resuming from unplanned hold situations. It is tied to information
in TMID, propagates constraints forward and backward to complete unspecified
values, and checks the plan against constraints. The advisor phase is a
situation advisor (TMSA) which proactively suggests tactics for resumption from
developed. A concept prototype for TMSA is under development. The TMID is
currently undergoing field testing. Displays for TMID and TMWI will be
described. Descriptions will include organization, rationale for
organization, implementation choices and constraints, and use by NTD.

Arthur E. Beller
NASA
Artificial Intelligence Section
DL-DSD-23
KSC, FL 32899
407/867-3224
FTS/823-3224

H. Greg Hadaller
Boeing Aerospace Operations
Advanced Computing Technology Group
FA-77
KSC, FL 32899
407/867-3770
FTS/823-3770

Mark J. Ricci
Boeing Aerospace Operations
Advanced Computing Technology Group
FA-77
KSC, FL 32899
407/867-3770
FTS/823-3770

Session H5: VIRTUAL REALITY

Session Chair: Dr. Steve Ellis

N92-22344

Visually Coupled Systems (VCS): The Virtual Panoramic Display (VPD) "System"

by

Dean F. Kocian
USAF Armstrong Aerospace Medical Research Laboratory
Wright-Patterson AFB, Ohio

Abstract

This paper describes the development and impact of new visually-coupled system (VCS) equipment designed to support engineering and human factors research in the military aircraft cockpit environment. VCS represents an advanced man-machine interface (MMI). Its potential to improve aircrew situational awareness seems enormous, but its superiority over the conventional cockpit MMI has not been established in a conclusive and rigorous fashion. What has been missing is a "systems" approach to technology advancement that is comprehensive enough to produce conclusive results concerning the operational viability of the VCS concept and verify any risk factors that might be involved with its general use in the cockpit. The advanced VCS configuration described here, has been ruggedized for use in military aircraft environments and was dubbed the Virtual Panoramic Display (VPD). It was designed to answer the VCS portion of the "systems" problem, and is implemented as a modular system whose performance can be tailored to specific application requirements. The overall system concept and the design of the two most important electronic subsystems that support the helmet-mounted components, a new militarized version of the magnetic helmet mounted sight and correspondingly similar helmet display electronics, are discussed in detail. Significant emphasis is given to illustrating how particular design features in the hardware improve overall system performance and support research activities.

Introduction

The last six years of advanced military equipment development activity, and particularly the last four, have been marked by a renewed interest in the application of visually coupled system (VCS) technology to the military aircraft cockpit. Not since the early to mid 1970s has the activity level been so intense for militarized versions of this technology. Much of the renewed developmental activity has, rightly, concentrated upon the helmet mounted display portion of the VCS. This fact is demonstrated by new helmet mounted displays (HMDs) from such manufacturers as Hughes Aircraft, Kaiser Electronics, GEC, and Honeywell, and from DOD activities like the Army LH Helicopter Program, the Air Force F-16 Night Attack Program, and the joint Navy-Air Force INIGHTS Program. The emphasis on HMD design, particularly its size and weight, is because they have been a major performance-limiting factor to the widespread safe use of the technology in ejection seat aircraft. However, the successful integration of VCS technology into the cockpit includes the solution to a number of utilization and performance problems that cross the boundaries of many technical disciplines (see references [01,02,07]).

It can be argued that, in many instances, a "best" approach to either investigating or solving VCS performance and utilization issues, is to have at hand state-of-the-art (SOA) hardware that has the required "programmability" to support engineering and human factors research and also the ruggedness to be operated in the final intended environment, in this case the military aircraft cockpit. Until the advent of the VPD development program, this type of hardware, representing SOA, dedicated instrumentation specific to the VCS technology area, had been missing. This paper investigates some recent development activities at the Armstrong Aerospace Medical Research Laboratory (AAMRL) involving the VCS electronics that directly support the helmet mounted components. The discussion covers the rationale surrounding their primary operational characteristics and their impact on the current state of VCS technology.

A "Research-Oriented" Visually Coupled System

Figure 1 depicts one variation of a VCS system, the Virtual Panoramic Display (VPD). This system is designed primarily to support VCS whose helmet-mounted displays (HMDs) use miniature cathode-ray-tube (CRT) image sources. The VPD provides the basic helmet tracking and display presentation capabilities. It also supplies the configuration programmability, interface flexibility, and self-contained data collection needed to support advanced research activities.

The display subsystem of the VPD, the HMD, may include either one CRT image source and one or two optical channels, providing a monocular/biocular display presentation, or 2 CRTs with dual independent optical channels making possible a true binocular presentation. The visual fields may be either fully or partially overlapped, and may be aligned, using programmable electronics, permitting, if desired, the presentation of stereoscopic images. The CRTs, employing, in most cases, narrow-band emission phosphors, may be any of the standard bipotential electron lens designs commonly available, or be of an advanced design including additional grid control
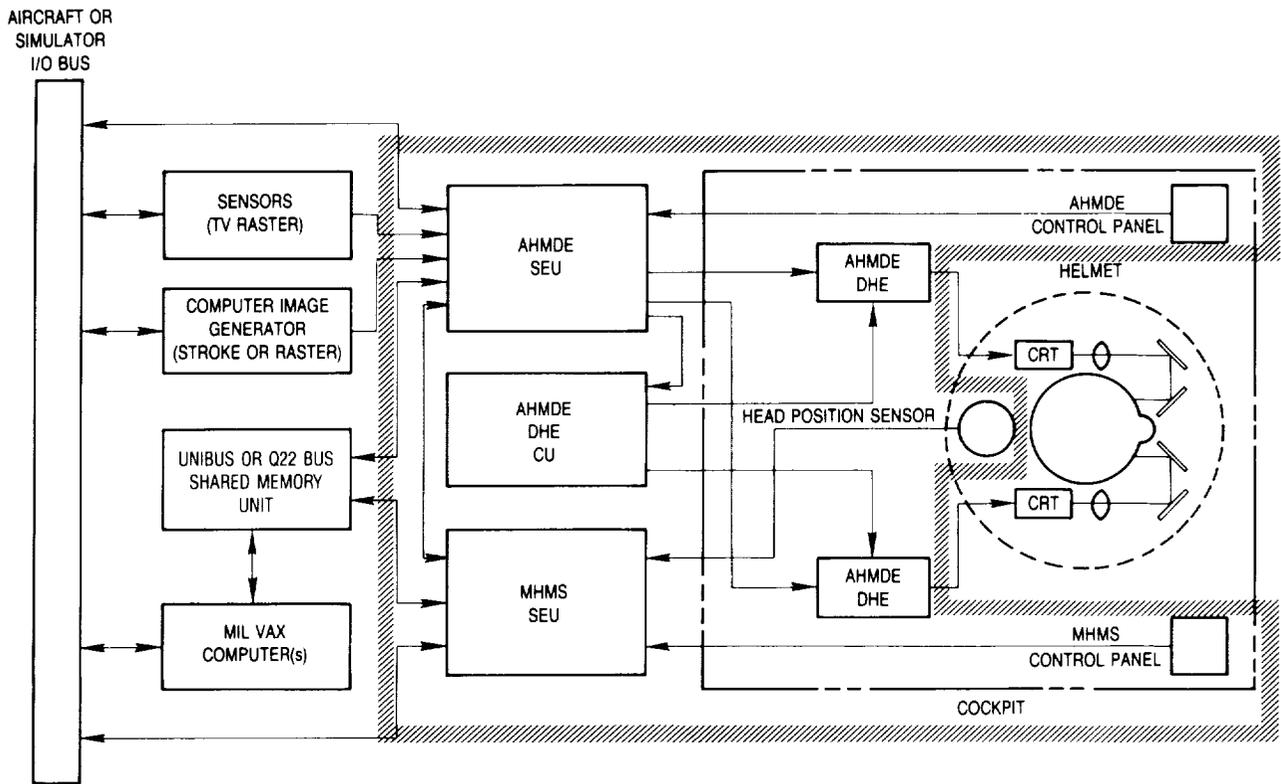
FIGURE 1
Block Diagram of Generic VCS System Hardware

elements. The CRTs are interfaced to specially-designed analog helmet-mounted display electronics (AHMDE) which "tailor" the displayed information to the requirements of both the HMD optics and CRT. A major thrust of current visual display research for the military cockpit assumes that mission equipment package (MEP) data and sensor-generated information must be "fused", in part visually, in some optimal manner on the HMD to help improve the pilot's moment-to-moment situation awareness. Thus, the VPD AHMDE has been designed to support a range of anticipated video input combinations for sensor-generated and computer-generated visual information that can be displayed simultaneously on the CRT.

Much of the displayed information must be changed and/or updated, based upon the pilot's instantaneous line-of-sight (LOS) and helmet position within the cockpit. To support this function, an AC magnetic helmet mounted sight (MHMS) is included to provide both helmet attitude and position vectors. The newer version of the MHMS used for this VCS configuration can be programmed to compensate for certain environmental disturbances to which it is susceptible.

To promote ease of programming and transition to-and-from ground-based and airborne research environments, a combination of Digital Equipment Corporation Q-Bus and UNIBUS processors and Motorola 68000 family VME processors have been

employed as both imbedded and stand-alone processors. To facilitate data transfer beyond the limitations of the military 1553B data bus, a variety of high speed interfaces including a multi-port shared memory (MPSM), which exists in both laboratory and militarized form, were developed. The MPSM allows up to ten DEC or VME-based processors to simultaneously perform parallel read and write operations between each other. To facilitate non-volatile memory storage, militarized hard disks are available for the DEC-based processors and EEPROM for the VME-based processors. This architecture fosters ease of expansion when additional processing power is needed, and permits additional enhancements, such as auditory localization and physiological test battery monitoring, to be added to the basic VCS, as needed and available.

The shaded area of Figure 1 represents the core components of the VCS electronics subsystem for most near-term VCS configurations. This paper will focus the remainder of the discussion on just the portion of the shaded region that includes the AHMDE and MHMS, emphasizing advancements that facilitate not only hardware performance, but also improve or facilitate the VCS interface and research activities. Some discussion in a similar vein relating to the helmet mounted display components can be found in [01,02].

549

The Helmet Mounted Display Electronics "The Key to Miniature CRT Performance"

When the VPD development program was begun there was no general purpose helmet mounted display display electronics (HMDE) capable of operating different types of HMDs with inputs from a variety of display input sources. The problem was further exacerbated by the ongoing performance improvements in miniature CRTs, which at that time, available off-the-shelf display electronics could not support. There were also a number of deficiencies relating to how disparate video images were processed and combined for display on the HMD. The initial approach attempted a comprehensive all-digital update of not only the display electronics functions but also graphics processing and image processing functions, that were felt to be necessary to support the complete visual-interface concept for a large or medium FOV virtual head-mounted panoramic display. When this proved too ambitious for available program resources, development work was refocused to concentrate upon just the badly needed primary display electronics improvements. Among the more important of these were;

1) Allow the display electronics to optimize each CRT's grid control voltages for best performance, and store these parameters in local memory for instant recall if a CRT change had to be made,

2) Permit complete system setup and adjustment, including HMD alignment pattern generation, and allow this to be accomplished from the cockpit without the need to pull boxes, adjust potentiometers, etc.,

3) Provide sufficient deflection, video, and high voltage power to support the latest advancements in miniature CRT technology,

4) Provide the best possible video and deflection signal quality because of its more noticeable visual effect under the magnification of the CRT format by the HMD optical system,

5) Provide two separate video channels for binocular HMDs, accepting two raster and one stroke video input to support all but the most sophisticated VCS applications,

6) Provide automated system input flexibility to permit the use of raster image sensors having a wide range of line rates, and allow the internal raster generator to be synchronized to external video raster sources, and

7) Accept the HMS signals directly, to provide the right combination of raster imagery rotation and translation for a specific HMD design at the refresh rate of the HMD electronics.

To appreciate the significance of the new AHMDE design, one must understand how the above features and functions contribute to VCS performance. To accomplish this, the design of the AHMDE display electronics must be described in the context of the performance requirements placed upon it by the HMD optical system and CRT image source.

Serious use of the HMD dictates that it be considered an application specific animal. Indeed, many applications of VCS technology have resulted in less than desirable outcomes because off-the-shelf HMD systems were utilized instead of a design intended specifically for a given application. The most general example of an HMD application is one involving a particular aircraft and sensor suite whose field-of-view and resolution performance are already defined. Normally, it is desirable to match the FOV of the HMD to that of the primary sensor, because this places the least demand on image source performance, allowing the best contrast, luminance and resolution conditions to be obtained from the CRT. It also permits a 1:1 ratio to be set between the display and see-through image. For a more completion discussion of these interface issues, see reference (02). The next step in the integration sequence is to insure that the display electronics can adequately tailor the CRT display format characteristics to support the given application. The AHMDE represents the first time that a militarized VCS display electronics has been designed from "scratch" to provide the type of system integration tailoring that is required.

The AHMDE: "Designing and Building a Better Display Electronics Mousetrap"

Figure 2 depicts the basic VPD AHMDE components. The heart of the system is the system electronics unit (SEU). This box contains all of the power supplies and computerized parameter control and signal processing circuitry necessary to support a wide range of separate signal inputs from stroke and raster video signal source inputs for two miniature CRTs and tailor these signals for proper output to most types of miniature CRTs now available or being developed.

The display head electronics (DHEs) contain the final video and deflection amplification circuitry, miniature militarized high-voltage power supplies, and a microprocessor to set voltage levels, gain, offset, etc. for each particular CRT's optimum performance characteristics. The deflection amplifier employs high-performance power FETs in a Class-A amplifier arrangement to maximize linearity and repeatability of pixel placement on the CRT. The video amplifier employs a low noise, low output impedance design to maintain signal bandwidth while driving a 6 to 8 foot helmet cable that may exhibit considerable distributed capacitance. There are two separate DHEs to permit more flexibility in the use of each video channel. Such a configuration allows either one binocular HMD or two separate biocular or monocular HMDs to be operated with one AHMDE SEU. To maximize power dissipation for even the most demanding experimental operating conditions and minimize DHE size, liquid cooling of the DHEs (distilled water or an appropriately conditioned liquid, if freezing temperatures are expected) is employed.

The requirement for liquid cooling necessitated the addition of a cooling unit (CU) box that incorporates a militarized pump, heat exchanger, and the regulated voltages for the high voltage power supplies located in the DHEs.

550
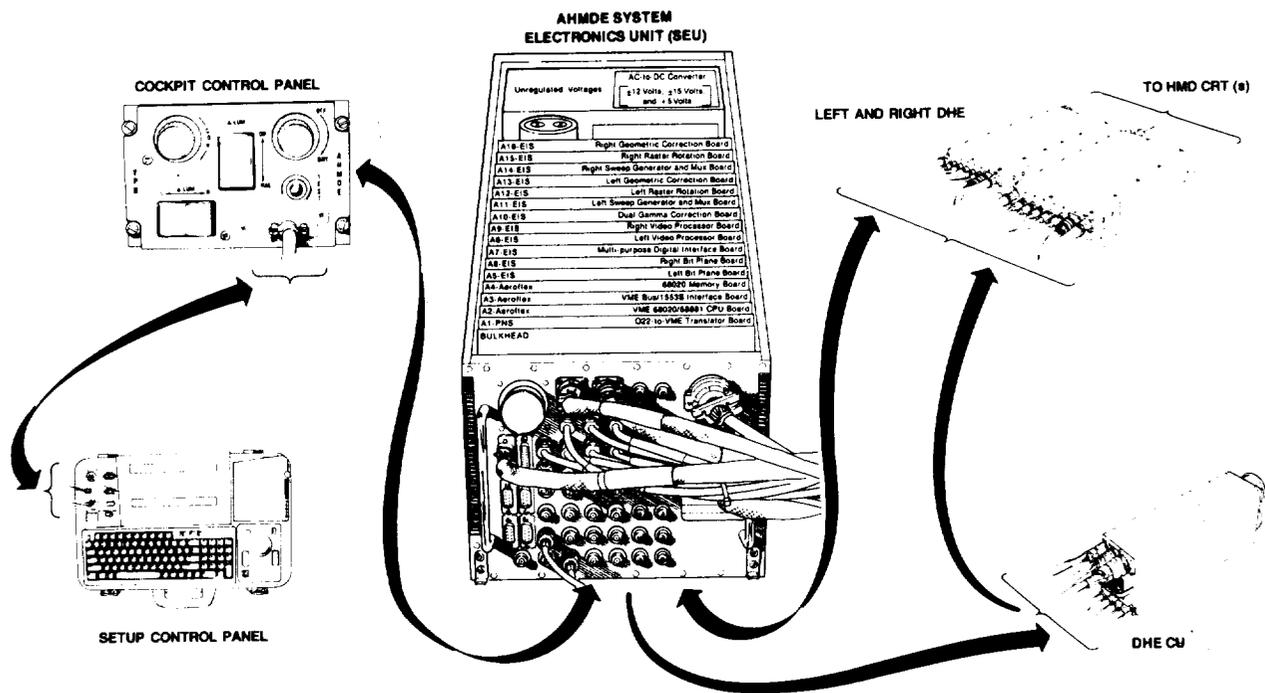
FIGURE 2
VPD AHMDE System Concept

A simple control panel (CP) that incorporates night-vision-goggle-compatible electroluminescent lighting allows adjustment of either CRT's stroke and raster video luminance and contrast. The CP also supplies the system interface for a unique setup control panel using a menu driven software interface that allows the experimenter or technician to adjust all critical system parameters, including HMD-CRT electronic alignment and CRT replacement, without removing the equipment from the simulator or flight test aircraft.

The AHMDE SEU

Figure 3 depicts a simplified block diagram of AHMDE SEU functions. The AHMDE SEU contains the bulk of the electronic signal processing circuitry to permit the AHMDE's raw signal processing performance and programmable features to be tailored for maximum benefit. The AHMDE SEU contains three primary signal processing groupings organized as digital and analog subsystems.

Digital Subsystem

The digital subsystem is comprised of the embedded VME processor and associated digital cards that implement data processing and transfer between external and internal destinations. Acting as the focal point for the AHMDE's extensive HMD video processing functions is a militarized 68020 VME CPU and processor board set. These boards store and execute the runtime code from a mixture of EPROM, EEPROM, and SRAM memory.

In order to support large, multiprocessor applications, an associated shared memory board provides parallel address and data transfers for VME and Digital Equipment Q22 or Unibus based computers. A 1553B serial interface for communication through the standard military aircraft digital bus and several RS-422 serial digital links are also provided. However, the proper use of all these external communication options requires the development of special software for application-dependent I/O drive routines to pass application-specific parameters from the aircraft or simulator bus to the AHMDE.

The AHMDE's internal memory also permits storing setup parameters for a number of HMDs and more than 100 different miniature CRTs. This allows the system to be adjusted for a variety of test configurations within a minimal time period. The AHMDE SEU, upon receiving the appropriate commands from either the cockpit or setup CP, selects the correct HMD alignment patterns from memory and moves the pattern data to the internal left and right bit-plane boards. These alignment patterns are normally configured for a specific HMD, and must be preprogrammed into the AHMDE's PROM memory. The bit-plane memory can also be used for the dynamic portrayal of simple space-stabilized head-up display (HUD) formats, if a particular VCS application does not have graphics processor electronics available. Again, as for the communication interfaces, the proper application-dependent software routines must be present to correlate the dynamics of such data with a particular symbol or alphanumeric character.

551

EXTERNAL VIDEO IN: EXTERNAL DEFLECTION IN:
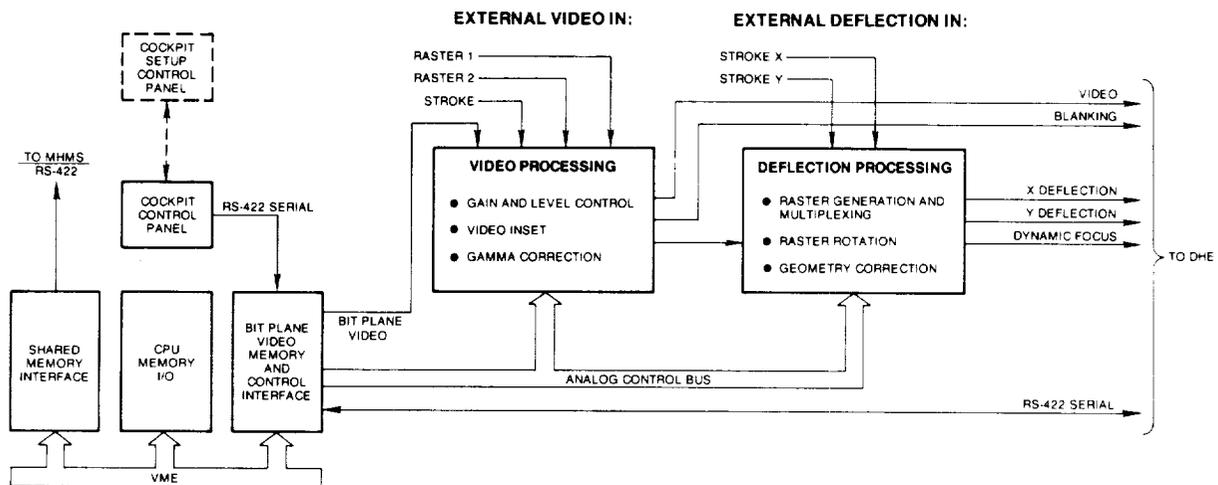
FIGURE 3
Simplified Diagram Of Major AHMDE SEU Functions

Analog Subsystem: Video Processing

The AHMDE's analog video processing section implements a number of important functions for the HMD visual environment. The normal gain and level controls have been expanded to separately accommodate two external raster sources, raster video from the internal bit plane image generator, and one stroke video source.

Raster 1 (see Figure 3) is designated the master raster and the internal bit plane video is synchronized to it, allowing it to be viewed separately, windowed, split-screened, or added into the raster 1 image. Although originally intended as a VPD HMDE function, the AHMDE does not digitally sample incoming video and perform scan conversions. Therefore, video on the raster 2 input must be in frame synchronization (FS) with raster 1. If only FS, but not line synchronization (LS), is present for raster 1 and 2, then they may be displayed in a vertical split-screen presentation. If both raster 1 and 2 inputs have the same FS and LS, a portion of raster 2 may be inset (windowed) into a user programmable portion of raster 1.

The AHMDE video processor also supports a number of more subtle requirements associated with VCS HMD applications. Each AHMDE video channel is designed to provide geometric shading and circular blanking (as received from the deflection processing subsystem). These functions compensate for video and optical effects associated with binocular HMDs (particularly those that are partially overlapped). They also support the often-present need to overscan the CRT raster in the horizontal direction to provide the largest CRT raster format possible for a given size CRT, operated at a specific aspect ratio (see reference [02]).

Many of today's high resolution, high luminance miniature CRTs often exhibit distinct nonlinearities in the luminance gray shade profiles. This condition is usually most noticeable with the brightest gray shades.

To allow the CRT luminance function to be corrected to a more linear and desirable luminance function, programmable gamma correction has been included.

Analog System: Deflection Processing

The special requirements of the VCS video environment have as much impact on the video deflection system, perhaps more so, as they do for the video signal functions. To support the video input options, each deflection channel incorporates two separate multi-rate raster generators. These permit switching between raster sources having different rates. When changing between rates, the AHMDE's sweep generators eliminate size adjustments with automatic size control.

VCS applications often require that the imagery displayed on the HMD be stabilized with respect to a coordinate frame other than the head as the head is moving about the cockpit or crew station. Depending upon the HMD configuration (see reference [02]), proper stabilization requires either a pure rotation or a combination of a rotation and translation of the display imagery as the head/helmet moves. The head tracking signals needed to perform the calculations for the imagery stabilization are obtained through a high speed serial bus connected between the AHMDE and a helmet orientation and position tracking system such as the VPD MHMS (as shown in Figure 1). The AHMDE supports this requirement only for raster based imagery. Stroke imagery must be rotated and translated at its source.

As explained in detail in reference [02], the CRT is a Tangent (theta) mapped system, while most HMDs obtain maximum optical performance and minimum weight using F(theta) mapped optical designs. Specific optical designs, whether they are monocular or binocular, may have other types of distortion that must be compensated for by nonlinear mapping of the CRT image. Finally, the CRT, itself, may have internal distortion requiring correction due to the interaction of

552

the deflection yoke, electron gun and e-beam. Given the magnitude and types of distortion associated with HMDs, even-order compensation, up to fourth order terms, would be desirable. However, signal-to-noise considerations in the AHMDE electronics effectively limit the geometric correction to third order terms. Thus approximations must sometimes be made for a given optical system, but the correction obtained is usually sufficient in subjective terms relating to viewer comfort, if not in terms of absolute measurements.

A larger CRT image requires less magnification by the optics. Since the exit pupil diameter for the HMD is the relay lens (more properly, objective lens) effective aperture divided by magnification, the relay lens can have a smaller diameter, i.e., a higher f-number. This makes it smaller, lighter, less expensive, and, possibly, of higher optical quality. To reduce the required optics magnification, and therefore working f-number of the HMD optics, the image is usually overscanned in the horizontal direction to enlarge the vertical height of the display format for a given aspect ratio. If this requirement and any rotations or translations force the scanned area to move off the phosphor quality area of the CRT, the AHMDE deflection circuitry incorporates circular blanking to extinguish the beam as it moves off the predetermined limits of the phosphor.

Taken together, the AHMDE's "designed in" functions allow it to drive almost any type of HMD now available or planned, using inputs from most standard video or stroke sources. Its programmability allows ease of use and repeatability of operating conditions for both the engineer and researcher. An itemized list of basic AHMDE performance is summarized in Table 1. A more complete description of the AHMDE's physical and electronic characteristics can be found in reference [06].

The Helmet Mounted Sight "Providing a More Complete System Concept"

Despite employing AC MHMS technology that has been available in some form for 20 years, the VPD MHMS really represents a more notable advancement in VCS technology than does the AHMDE. Many critical pilot activities involve, to some degree, the rapid acquisition of information, the accurate and fast positioning of display symbology depicting system state, and, after suitable cognition time on the pilot's part, the execution of one or more aircraft system state changes. If the man-machine interface (MMI) through which this interaction is being effected is a VCS, then the spatial relationships and positional accuracies of information portrayed on the HMD, as determined by the HMS position and orientation (P&O) tracking data, are extremely important. Not only must the quality of HMS attitude and position information meet some known and repeatable baseline level of accuracy, but it is also desirable to enhance the VCS's immunity to environmental disturbances to which the HMS or human operator are susceptible (at least in a "signal-to-noise-sense" to a threshold near or slightly beyond the limits of system-aided human perception). The magnetic helmet mounted sight (MHMS), despite the complexities of correcting for electromagnetic scattering, is still regarded as the HMS system of choice because of its rugged small transducers, immunity to other types of environmental problems associated with military vehicles, and the speed and accuracy of the six-degree-of-freedom (6DOF) orientation and position data that the MHMS can provide.

When the VPD development program was begun, there were still a number of significant deficiencies that had been noted concerning the operational characteristics of the AC MHMS. Among the more important of these were;

TABLE 1
Summary of Basic AHMDE Performance

Item	Parameter	Performance
1	Video line rates	Any line rate to 65 kHz scan rate
2	On-axis linearity	At least 0.25%
3	Spot motion and jitter	<0.0005"
4	Step response Small signal (10% display width)	<800 nsec to 0.1% of final position [Rise time = 816 nsec, Fall time = 970 nsec to 0.1%]
	Large signal (100% display width	<2.5 μsec to 0.1% of final position [Rise time = 2.3 μsec, Fall time = 2.39 μsec to 0.1% at 4.6 amps peak-to-peak]

Item	Parameter	Performance
5	Z-axis bandwidth	60 MHz (−3dB) [62 MHz (−3dB)]
6	CRT grid voltage stability Anode Ripple Regulation Other control voltages (G2, etc.) Ripple Regulation	Better than 0.05% Better than 0.2% Better than 0.05% Better than 0.1%
7	Distortion compensation	To 3rd order with cross-product terms
8	Delay lines	10–630 nsec @ 10 nsec steps

[ ] Indicates actual measured system performance

1) Its susceptibility to radiated electromagnetic interference from the head-mounted miniature CRT magnetic deflection yoke(s),

2) Limitations in the motion box of the helmet sensor,

3) Line-of-sight (LOS) accuracies that were not compatible with the HUD accuracies,

4) Update rates that were marginal for some applications, and could not usually be "synced" to a master system clock, and

5) Its inability to support research into biodynamic interference suppression.

Figure 4 depicts the system concept, organization, and primary system elements of the new VPD MHMS.

The key to improving the performance of the advanced VPD MHMS system is its new position and orientation (P&O) algorithms developed by Green Mountain Radio Research Company (GMRR) under AAMRL's direction, and the hardware parallel-

processor architecture developed by the Kaiser Electronics subsidiary, Polhemus, Inc (PI). The new algorithms improve system tracking accuracy and permit the effect of many environmental disturbances to be substantially reduced. The primary P&O tracking algorithm is a minimum variance linear estimator (MVLE). The MVLE makes direct use of the magnetic-field characteristics, and satisfies three objectives: (1) obtaining a least-squares best fit to the measured magnetic fields, (2) providing minimum expected mean-square error in P&O, and (3) providing maximum-likelihood P&O estimates. Under conditions of very rapid and continuous head movements, the MVLE algorithm cannot provide convergence to an accurate P&O solution. As shown in Figure 4, 'supervisory' software checks for such conditions, and, when detected, switches the P&O solution process to a nonlinear estimator (NLE) algorithm. The NLE makes direct, noniterative estimates of P&O, and, while its accuracy is not as great as the MVLE, its stability is absolute. Thus the MVLE and NLE are complementary. Under normal operating conditions, the MVLE provides the most accurate P&O estimates, while the NLE ensures correct initialization and recovery from very large and continuous step inputs
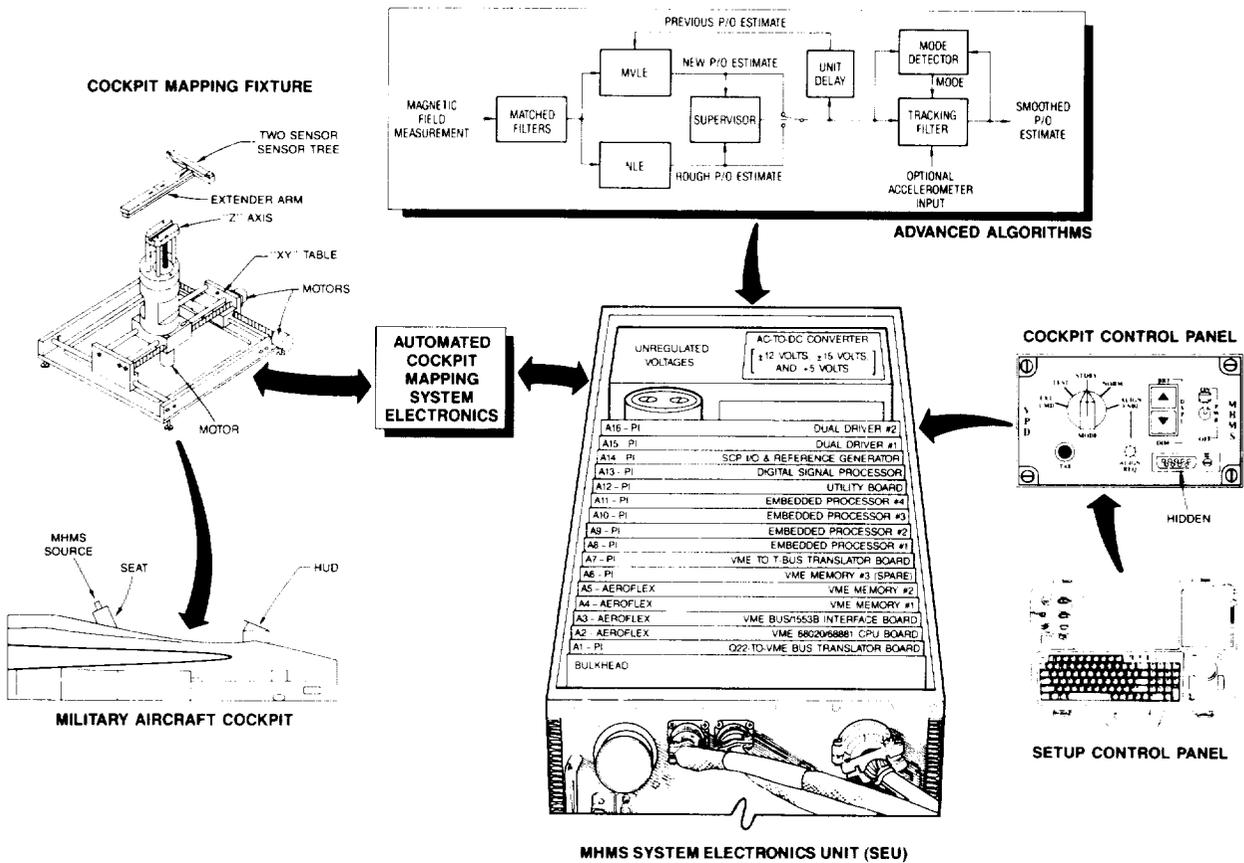
FIGURE 4
VPD MHMS System Concept

554

and power disruptions. Signal oversampling and a "matched filter" are used on the front end to obtain the needed signal-to-noise ratio (SNR) and to minimize head-mounted CRT interference. The hardware provides the necessary computational power and system upgrade flexibility, including optional accelerometer inputs. This flexibility permits adding tracking filter implementations to the MHMS outputs, whose programmed characteristics allow selective modification of its outputs, based upon environmental disturbances and the operator's tracking inputs. The primary purpose of the filtering is to improve operator LOS placement accuracy and to improve information extraction from the HMD.

The primary update rate for P&O tracking of a single helmet is an impressive 240 updates per second (UPS). Figure 5 indicates that as many as 2 observer heads and 4 hands can be independently tracked using separate P&O sensors. However, this maximum configuration reduces the update rate to 60 and 30 UPS for the heads and hands, respectively. Separate update cycles are needed to first sample and filter the raw field data, and then determine true P&O estimates. This reduces the data throughput rate to at least half of the maximum update rate. As shown in Figure 5, the MHMS can be "synched" to an external source or clock. This feature can be important for applications involving computer-generated imagery where obtaining equal update rates of head P&O is

desirable for the display of moving objects on the HMD. A set-up control panel (SCP) can also be interfaced directly to the cockpit side of the cockpit control panel (CCP). The SCP allows built-in test functions to be accessed and system parameters modified without removing the system from the aircraft.

The MHMS system is complemented by a programmable, semi-automated mapping fixture which, during the cockpit mapping and compensation process, is connected to the actual MHMS system that will be installed in the cockpit. The mapper's computer shares data with the MHMS processor hardware. Using a predetermined system error budget, the mapper system gathers preliminary raw magnetic field data based upon a completely quantitative process for allocation of the field mapping data points. Thus, the density of the sampled cockpit field data is controlled by solutions derived directly from computations involving the actual MHMS hardware and MVLE/NLE algorithms. A similar approach is used for mapping the magnetic field conditions induced by helmet-mounted scatterers, with respect to their fixed relationship to the helmet-mounted sensor, and, then, computing the needed field compensation coefficients. Taken together, this new "systems" approach to the MHMS provides superior tracking accuracy and complements the overall quality of the VCS MMI.
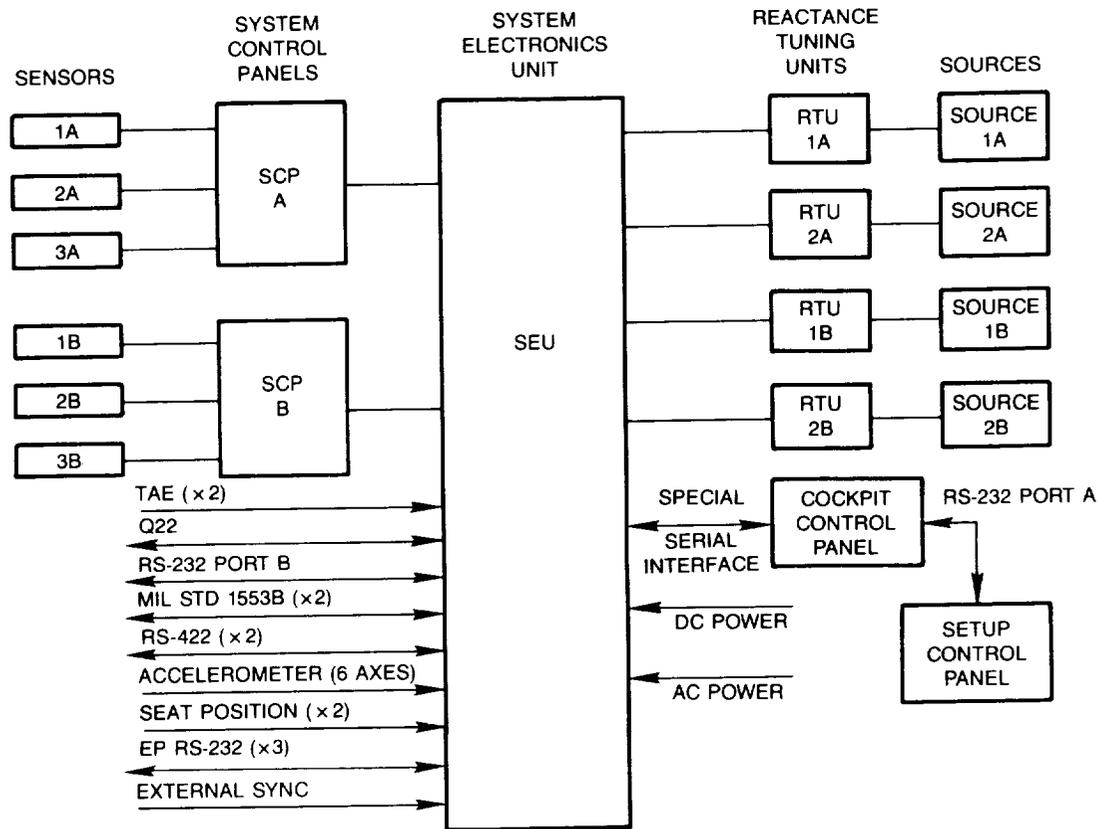

FIGURE 5
VPD MHMS Component Block Diagram

555

MHMS Performance: "Technology and Research-Based Requirements"

At the beginning of the VPD MHMS program, a set of performance goals was formulated based upon a programmatic requirement to demonstrate useful technology advancement and, also, the need to support research in operational aircraft environments. The most important of these were:

1. Enlarge sensor motion box to provide reliable coverage of head and hand position and orientation throughout the cockpit volume,

2. Improve static accuracy to levels approaching that of the HUD (1 to 3 milliradians) to support most weapon system interface functions,

3. Improve update rate and throughput rate to support auditory localization, as well as visual subsystem data requirements,

4. Provide a system capable of investigating techniques and equipment that enhances system and human operator immunity to external disturbances, thus, allowing the improvement or modification of VCS signal processing functions, and

5. Implement a system supporting research oriented activities that could be used in both ground-based and airborne environments.

System Concept Basis: "Establishing System Error Criteria"

Once system requirements were set down, it quickly became apparent that a new approach to the MHMS software and hardware was required. For the hardware, it was necessary to develop an affordable militarized architecture that was flexible and offered high computational rates. For the algorithms, whose characteristics would largely determine overall performance, a joint program with GMRR was initiated to identify how the data in the MHMS fields could be used to compute P&O with minimum error, and identify factors that limited the computational process (see references [04,05]). For the MHMS system, allocation of a total system error budget was made early, and a concerted effort was made to meet it, particularly because of the parallel development process used in the program.

The system error budget and design equations are based upon two basic concepts: (1) expected mean-square measurement error (EMSME) is convertible into an expected mean-square estimation error (EMSEE) through a scale factor and (2) EMSME is expressible as the root mean square (RMS) of the various error sources. Reference [03] contains a more complete development as to how this error performance was determined.

Improving MHMS System Accuracy: "Static" and "Dynamic"

The desired error performance is demanding and exceeds the performance of the standard 12-bit MHMS systems now available. The key factors helping the new system to meet this ambitious goal are:

1. Improved MHMS algorithms that embody associated analysis explicitly defining and characterizing system error sources,

2. An integrated and automated mapping fixture, and

3. MHMS hardware that includes special front-end digital signal processing hardware, improved update rate and resolution, and a slightly larger source (radiator or transmitter) size.

An adequate system solution has required a more technically-complete approach to the MHMS development. As a result of the early and in-depth application of applied mathematics, the minimum variance linear estimator (MVLE) algorithm was evolved and has become a key element for improved MHMS performance. The MVLE has two especially desirable properties when used with a magnetic P&O tracking system [04]: (1) it provides the most accurate estimates from noisy measurements and (2) rather than assume a free space condition that must be corrected, as other MHMS algorithms have done, it makes direct use of the magnetic-field characteristics at the sensor. It is the second property of producing a correct estimate directly from the magnetic-field conditions that allows the MHMS sensor to be tracked more accurately, even down to conducting metal surfaces. The algorithm design also enhances the incorporation of moveable scatterer compensation (for the head mounted CRTs) into the primary P&O algorithm. Oversampling of the MHMS signals and a matched-filter optimize reception of the separate and simultaneous three-axis winding signal excitation frequencies. The matched-filter can also provide attenuation of the radiated CRT deflection noise at the line rate frequency for the HMD raster imagery. The matched-filter implementation is essential to ensuring that an ~70 dB signal/noise ratio (SNR) for worst case conditions is met. The SNR is also improved by a slightly larger source, true 14-bit resolution of the magnetic fields, and distance-related gain changes in the radiated B-field strength. A whole host of signal processing refinements, beyond the scope of this paper, are used to improve P&O tracking performance, including compensation for finite transducer size effects, seat-movement compensation, and reductions in the buildup of computational errors. Because the MVLE tracks incremental changes in the magnetic fields, a higher update rate improves its stability. The operation of the MVLE is signal-dependent. It can remain stable for large step inputs of several hundred degrees/second, if followed by relatively static field changes or if fed continuous incremental inputs per update cycle that change by no more than about one to two degrees. Given the new MHMS's update rate (up to 240 updates/second), the MVLE can remain stable for head movement rates of hundreds of degrees per second. However, the performance boundary is made fuzzy because of its dependence on changes between incremental updates that are determined by head movement and system configuration-dependent update rates. Therefore, as Figure 4 shows, the software includes a supervisory process which switches the MHMS P&O solution to a NLE whenever stable solution criteria are not met. A sample of simulated MVLE tracking performance is presented in Figures 6a and 6b. The figures clearly show that the presence of fixed scatterers affects the minimum error that can be obtained, but the error
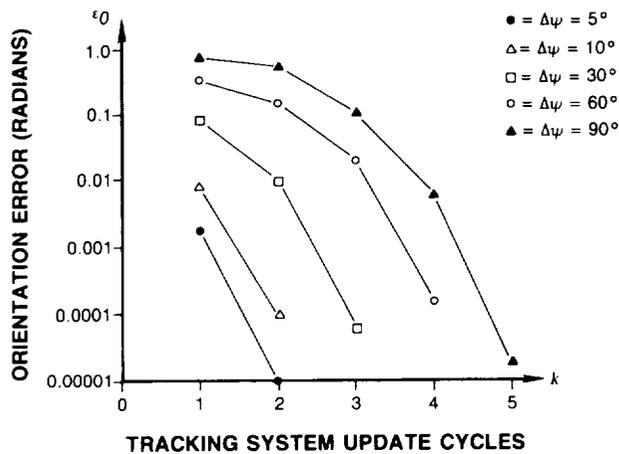
6a) Free Space Error Performance
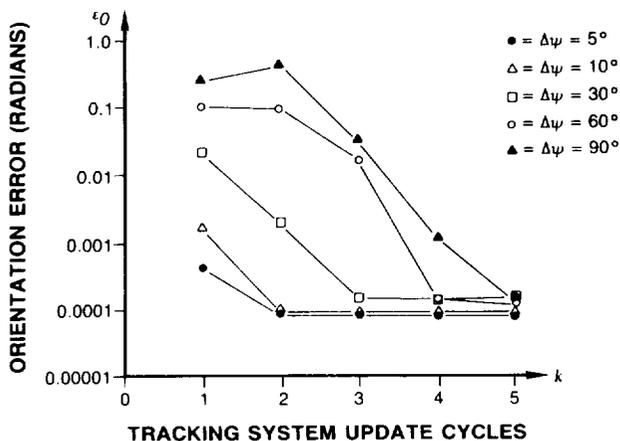
6b) Error Performance in the
Presence of 2 Fixed Scatterers

FIGURE 6
Representative MVLE Orientation Tracking
Performance for Free-Space and
Fixed Scattering Environments

floor (0.0001 radians or 0.34 arc minutes) meets the system error requirements listed in Table 2.

Table 2 also depicts important VPD MHMS performance parameters including the expected static accuracy of the new system when the automated mapping fixture is used to determine the allocation of field measurement points from which the necessary system compensation coefficients are computed.

MHMS Integration Issues

P&O Update Rate and Resolution

The implications of the new MHMS performance for the VCS are important for both operational and test applications. Integration issues are also more subtle and dependent on the characteristics of the MHMS technology than the AHMDE, which generally adheres to standard video practices. The new system algorithms and hardware improve static accuracy to an estimated 1 or 2 milliradians within the "HUD Box" (± 30° in azimuth and elevation), and to about 4-6 milliradians throughout the entire sensor motion box. Accuracy is also quantified better by the interactive mapping fixture design, which is an important improvement for research oriented activities. Resolution has also been improved to better than 0.4 milliradians. Resolution can be a significant parameter for head-driven display presentations where small head movements can be detected on the HMD presentation under conditions of high apparent magnification. 12-bit P&O tracking systems have been observed to cause undesirable and detectable discrete jumps in the location of the display imagery on the HMD during small head movements. 14-bit systems seem to provide enough additional resolution to make this artifact virtually undetectable. There are tradeoffs associated with the improved performance. Achieving the added SNR needed to attain an honest 14-bit system requires a larger source. The larger sources measure 1.25 inches to 1.5 inches square and weigh between 7.5 and 9 ounces. The ideal mounting location in fighter aircraft is on the cockpit canopy behind the pilot. The larger and heavier MHMS source may be too heavy for mounting in some cockpit canopies (e.g. the F-16) because birdstrike induced mechanical canopy waves are more prone to cause canopy failure with the heavier MHMS source mounted in them.

Most MHMS system P&O algorithms require at least two update cycles to obtain good convergence to accurate measured dynamic head location outputs. The new MHMS is no exception, as Figure 6 data indicates. The higher update rate reduces the convergence latency problem down to manageable levels. It also aids the throughput delay problems for computer-generated imagery systems which must place their imagery on the HMD according to the MHMS P&O updates. For applications, such as auditory localization, even higher update rates may be desired, because the ear can follow position update latencies in sound field vectors of less than one millisecond. A higher update rate also aids one area of MHMS performance that is particularly hard to quantify: system dynamic accuracy. The problem with this requirement is its measurement. Past development efforts that have investigated this problem have resulted in budgetary estimates of 3 to 4 hundred thousand dollars to produce an adequate test fixture. This is an amount that meager development budgets have not been able to handle with competing commitments of greater overall import. Perhaps a good alternative for the HMS is the achievement of higher update rates which reduce the latency between the measured and real head position and orientation and, thus, inherently improve system dynamic accuracy.

557

TABLE 2
Summary of Basic MHMS Performance

Item	Parameter	Performance
1	Angular coverage	
	Azimuth	±180°
	Elevation	± 90°
	Roll	±180°
2	Normal cockpit motion box coverage (can be essentially the physical limits of the cockpit)	
	X direction	±21″
	Y direction	±15″
	Z direction	± 8″
3	Static angular accuracy [line-of-sight (LOS) error for reduced angular coverage of ±70° in azimuth/elevation and ±30° roll]	
	50% circular error (CE)	<0.2° (2–3 milliradians)
	99% CE	<0.4° (6–7 milliradians)
4	Static angular accuracy (for conditions stated in 1 above)	
	50% CE	<0.3° (4–5 milliradians)
	99% CE	<0.5° (8–9 milliradians)

Item	Parameter	Performance
5	Static translational accuracy	
	X, Y, and Z	≤0.2″ plus 1% of separation distance between source and sensor
6	Resolution	
	Angular	~0.022°
	Translational	Better than 0.05″
7	Repeatability	Twice resolution limits in (6)
8	Update rate(s)	
	a) MHMS with basic tracking algorithms	240 updates/sec (U/S)
	b) Two-cockpit operation	120 (U/S)
	c) With virtual hand controller	
	1 head/1 hand	120/120 U/S
	1 head/2 hands	120/60 U/S
	2 heads/1–4 hands	60/30 U/S

System Design and the Operating Environment: Improving Pilot Confidence in VCS

Using a helmet orientation and position tracking system to direct weapons and place/stabilize imagery on the HMD, makes consistent system P&O outputs a necessity. Obtaining P&O tracking consistency for the MHMS involves two major system integration issues.

The first is where the pilot must or might place his helmet during any mission eventuality. For most VCS applications, it is desirable to have the entire head movement range under which functional display presentations are to be maintained, covered by normal MHMS operation. This viewpoint is often confirmed by experience with test pilots and the feedback that they provide concerning their experiences with VCS. This circumstance can occur in many present MHMS systems when radiator-sensor range is exceeded or unusual pilot head attitudes occur during mission activities which place the helmet mounted sensor near electrically conductive surfaces. Such artifacts are deemed to be unacceptable by most operational personnel with extensive VCS experience. As mentioned, the cockpit mapping and system compensation approach employs an automated mapping fixture. The computerized control interface between the mapper and MHMS system and its algorithms permits precise sampling of the cockpit electromagnetic field environment to ascertain that the desired error performance, cockpit environment, the system software, and operational helmet configuration are adequate. Until the development of the new MHMS, movement of the helmet sensor very near to conducting surfaces could produce wildly jumping display imagery. An undesirable solution to such problems has been to freeze LOS signals at their last known "good field" condition, but the HMD imagery as positioned by the HMS will not reflect true P&O. The

new MHMS algorithms now permit the sensor P&O to be tracked down to conducting metal surfaces and permit the graceful degradation of system accuracy and resolution to 13-bits, 12-bits, etc. for conditions where the maximum full resolution source-sensor range is exceeded. Therefore, a more stringent sensor tracking requirement that favors reliability of the MHMS-controlled placement of the HMD scene contents does not now imply a new, high risk development effort. It does imply, however, a greater allocation of computational power to the MHMS function than is offered by other available variations of the MHMS, militarized or commercial.

The second major integration issue is the spatial relationships of both the other helmet components (especially those that can distort or attenuate the MHMS magnetic field) and the MHMS source (transmitter) with respect to the helmet sensor location. The major, but not necessarily the only, significant helmet-mounted scatterer is the CRT. Figure 7 illustrates the nominal location volumes of the helmet mounted CRTs for the usual binocular and monocular/biocular HMD configurations with respect to the MHMS sensor. The ideal mounting location for the MHMS is on the crown of the helmet, in part, because this location is distanced from many sources of interference, and because the ideal location for the MHMS source is to the rear of the pilots head in the cockpit canopy or above the seat back. CRT locations 1 and 2, shown in Figure 7, can be particularly bothersome to reliable MHMS operation for head attitudes that include a directed LOS toward the side or rear of the aircraft with positive elevation angles. VCS integration planning should give consideration to such relationships and the tradeoffs they imply. In so doing, an attempt should be made to place the sensor on the helmet and the transmitter in the cockpit at positions where obscuration of the transmitters signal (by a helmet CRT, etc.) due to head movement is unlikely or

558

COCKPIT CANOPY

MHMS SENSOR

MHMS SOURCE

```
┌─ ─ ─ ─┐
│       │ — INDICATES MOST LIKELY LOCATIONS FOR HMD CRT
└─ ─ ─ ─┘
```

NOTE: POSITION 1 FOR CRT LOCATIONS SHOWN IN TOP VIEW ONLY FOR CLARITY
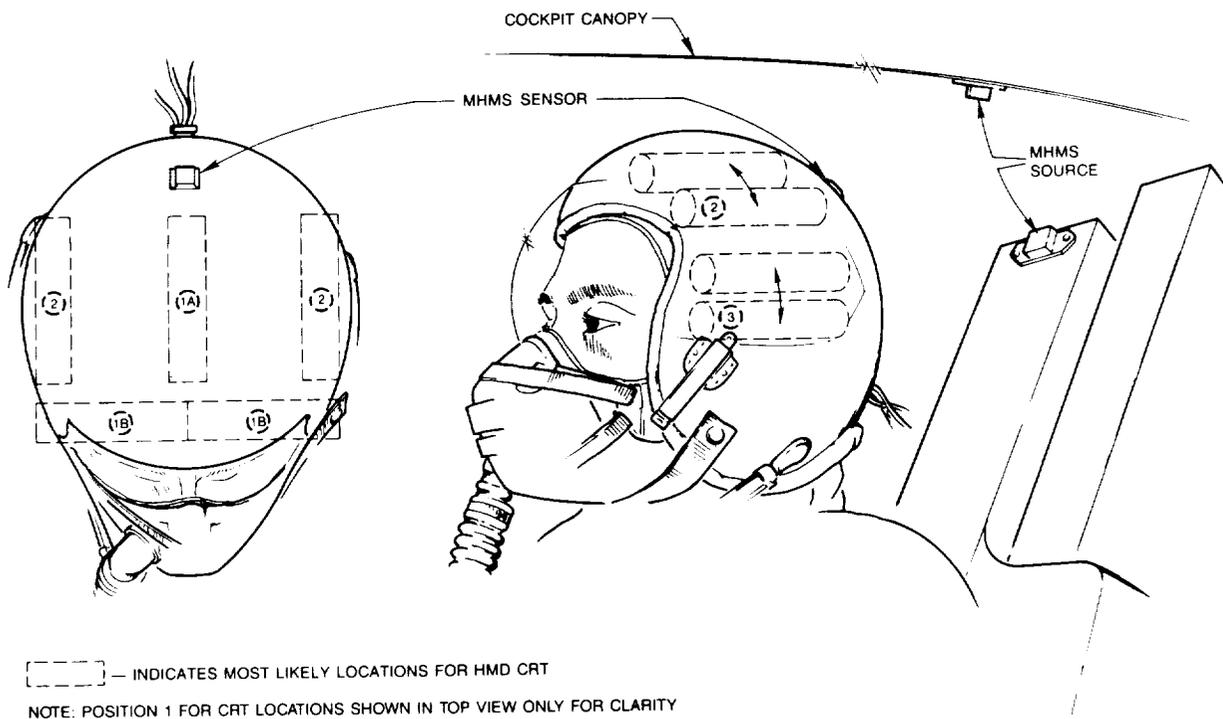
FIGURE 7
Normal Position Envelopes for Helmet-Mounted CRTs

impossible. Dynamic compensation for the movement of the head mounted CRT(s) may be critical to overall system tracking performance, especially if two CRTs are used or the mounting geometries of the CRT(s) and sensor cannot have optimum locations. A top-mounted CRT centered on the crown of the flight helmet for biocular HMD designs often poses the most serious problem.

One former approach to compensating for dynamic scatterers was to characterize the secondary field of the moveable scatterer and compensate for its effects using a dipole or multi-dipole model. This technique has not worked well, because the dipole locations are not readily determined and their scattering parameters depend on the number of dipoles used (quickly raising computational overhead) and an accurate knowledge of the location of each dipole. Recent experience with an alternate approach (05) that characterizes the scattered field as a sum of the multipole fields appears to have produced superior results in laboratory testing. In this technique, the multipole moments are linearly related to the primary magnetic field and its gradients by a set of scattering coefficients that are readily determined from sets of scattered-field measurements by linear-coefficient fitting techniques.

The effect of fixed and dynamic scatterers on MHMS operation is important enough in the design and integration of VCS that a short discussion of basic considerations is appropriate. An initial assessment of

the MHMS environment helps reduce potential problems during final system integration by noting a few geometric relationships and the worst-case conditions for MHMS field distortion caused by fixed and helmet-mounted moveable scatterers. Figure 8 depicts the worst-case situation for a fixed scatterer in relation to the transmitter(source)-receiver(sensor) distance. This occurs when the sensor is directly below the source. Since amplitude for the MHMS quasi-static field is proportional to the inverse cube of distance, the worst case distortion ratio (D) is given by equation (1) where "p" represents distance. Using (1) and a ratio of 5.5 for p_2/p_1, one computes a distortion ratio of $1/(11 - 1)^3 = 0.001$ or 0.1 percent. A ratio of 2.82 would produce a distortion ratio of 1 percent.

$$D = (p_1)^3/(2p_2 - p_1)^3 = 1/(2p_2/p_1 - 1)^3 \quad (1)$$

Where: $p_3 = 2p_2 - p_1$

Work performed during the development of the VPD MHMS at AAMRL and GMRR and documented in (05) helped to characterize the field distortion due to the presence of several types of simply-shaped dynamic scatterers. These include spheres and prolate spheroids which resemble the shape of the normal cylindrical CRT. For the moveable scatterer, the ratio between scatterer-to-sensor-distance (p') and scatterer

559
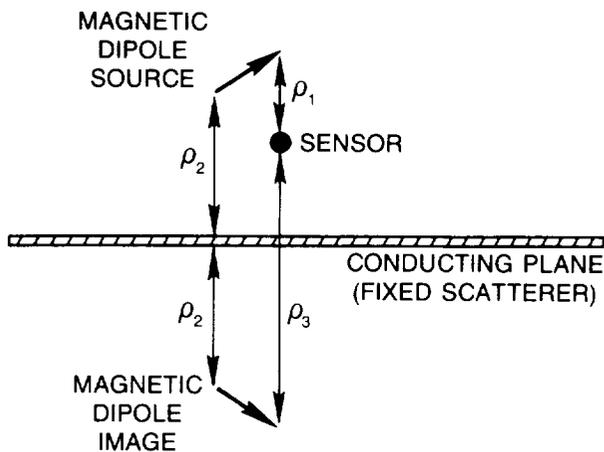
FIGURE 8
Fixed Scatterer Relationships

size (L) becomes the dominant relationship for computing the distortion ratio. For the inverse cube relationship we now get the relationship shown in equation 2. As shown for the spherical scatterer in

$$D = (L / \rho')^3 \qquad (2)$$

Figure 9, the distortion contours are simply spheres centered about the scattering sphere. For example, in Figure 9, a length of 4.64 L (where L = radius of the scattering sphere) produces a distortion ratio of 1 percent. As compared to a sphere with radius of L_r, a prolate spheroid with semiaxis of $L_1 = L_r$ and $L_2 = L_3$ = 0.2 L_r, representing about the eccentricity of the HMD CRT, produces maximum scattering no worse than about twice that of the sphere. The point being
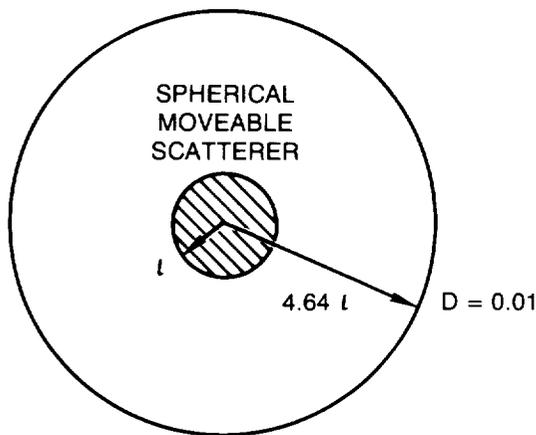

FIGURE 9
Moveable Scatterer Distortion
Contours Near A Conducting Sphere

made is that the helmet system design greatly affects the method and complexity of the required dynamic compensation process in the MHMS. If the HMD design permits a lower mounting location for the CRTs, as shown for location 3, in Figure 7, then the MHMS sensor may be placed at or near the top of the helmet. Located in this manner, the CRT scatterer(s) have a much reduced probability of blocking the MHMS signal during head movement. The minimum scatterer/sensor distance ratio will also be relatively large. A biocular design, as mentioned earlier, may, with most preferred sensor mounting locations, provide an increased opportunity for the scatterer to block the source's signal to the sensor. Alternately, placing the MHMS sensor on top of the helmet near the CRT scatterer to prevent signal blockage by the scatterer will significantly reduce (in a negative manner) the scatterer-to-sensor-distance to scatterer size ratio.

It is also wise, and usually necessary, to imbed some sort of capability into the MHMS to overcome situations where the HMD CRTs will completely block the signal to the receiver until an unobstructed line-of-sight (LOS) can be reestablished with the transmitter. The only good alternative is usually to hold the MHMS P&O output data at the last good field data point until radiated field conditions return to acceptable levels.

MHMS system signal properties are also influenced significantly by the material composition and structure of the overall VCS design. Of greatest concern are materials which, because of their physical properties, severely attenuate or distort the primary magnetic fields. For example, conducting metal on the helmet, per se, does not significantly degrade the uncompensated performance of the MHMS. Rather large conducting surfaces, structures that allow magnetic field induced current loops to form, and ferromagnetic materials are the major contributors to reduced performance (i.e. greater system error). In particular, the VCS system integrator should examine its subcontractor's production techniques to ensure that unexpected problem sources will not be present. One significant problem involved aluminum shimming rings for the HMD relay lens elements which were discovered to be causing severe dynamic scattering of the MHMS fields that was hard to compensate to obtain satisfactory error performance. The solution was to place a notch in each ring to prevent formation of current loops and secondary magnetic fields. Thus, an improved system design, as represented by the VPD MHMS, and thoughtful integration can complement each other to produce a more reliable pilot-centered or operator-centered system.

Optimizing Sight and Display Utilization

In this author's opinion, optimization is always a relative concept - for one usually does the best that can be done at a given time with available resources. Among the VCS optimization issues that have resisted a major improvement are the degradation of HMD performance caused by aircraft vibration transmitted to the head/helmet, and the effects of system P&O delays or head movement artifacts on operator LOS tracking. LOS tracking can be further divided into perhaps at least seven categories or modes, which may

560

have differing implementations for optimal tracking filter solutions. A possible set is:

1) Pointing (at a static target),
2) Tracking (a distant moving target),
3) Close Tracking (target moving close to an observer),
4) Handoff (of LOS from one observer to another for any mode 1-3),
5) Searching (for a new target),
6) Transition (between one LOS and another), and
7) Wandering (no specific LOS objective).

Both aircraft vibration and system induced LOS transport delays for display symbology positioning can severely degrade the operator's ability to extract and/or use information presented on the HMD. As shown in Figure 4, the MHMS electronics unit contains four imbedded processor boards. One of these processor boards was intended for use in aiding the investigation of the biodynamic interference suppression (vibration) and LOS tracking issues. Figure 5 depicts accelerometer inputs that can be used as an option for receiving direct inputs of head and aircraft vibration. The use of accelerometer inputs can provide improved stabilization of the HMD image, and needs further investigation. Often, however, systems utilizing accelerometer inputs exclusively perform poorly during large rotations of the head. Stabilization of the LOS signals, as derived from the MHMS P&O updates, using adaptive filtering, is also possible. The effectiveness of such stabilization may be reduced, though, by the possible relatively large phase errors that can occur in attempting to stabilize over a 5-Hz bandwidth with samples taken at 30 or 60 Hz. The improved update rate of the VPD MHMS and the use of Kalman filters or complementary filters (actually a form of Kalman filtering) that combine measurements of head P&O with accelerometer outputs may overcome these difficulties. The VPD MHMS provides not only the capability to run such algorithms, as they could be run in any computer simulation environment, but also the capability to test them in relation to operator performance in the actual airborne environment.

The programmable flexibility of the VPD MHMS also allows the benefits of advanced cueing modes to be implemented and studied. One example is coordinate intersection cueing (CIC). In a CIC mode, physical cockpit switches, or switches imaged onto a panel mounted display, are referenced to the MHMS source's coordinate system. Utilizing the 6 DOF measurements of the MHMS, the location of these switches is constantly recomputed by the MHMS, allowing "no-hands-needed" LOS activation by the pilot with his MHMS LOS reticle. This mode effectively duplicates a portion of the oculometer function without the need to add this hardware to the helmet.

Summary

The VCS MMI represents a significant departure from the standard cockpit MMI in use today. The development approach for the major VPD systems comprising the VCS MMI, should, if properly utilized, aid the transition to this advanced cockpit interface. For both the AHMDE and MHMS, the closed-loop system implementation for installation and operation should aid both the researcher and manufacturer. The

benefits for research personnel would be extremely reliable data describing the accuracy of the VCS implementation and very flexible airborne-qualified test instrumentation. The result for a potential manufacturer should be better a priori knowledge concerning interface and production issues. For example, a calibration process that is reliable and accurate enough to guarantee that the mapping process for one cockpit can apply in a labor savings manner to like configured cockpits, thus permitting the mapping of one cockpit to suffice for an entire block of similar aircraft. Whether this system approach meets expectations remains for the delivery of the production systems in 1991 and out-year testing and experience to confirm.

References and Selected Bibliography

[01] Buchroeder, R.A. and Kocian, D. F., Display System Analysis for the LHX Helicopter Application, US Air Force Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB OH, 1989, AAMRL-TR-89-001.

[02] Kocian, D.F., Design Considerations for Virtual Panoramic Display (VPD) Helmet Systems, AGARD Symposium on The Man-Machine Interface in Tactical Aircraft Design and Combat Automation, AGARD-CP-425, July, 1988.

[03] Kocian, D. F. Visually Coupled Systems (VCS): Preparing the Engineering Research Framework, IEEE, AESS, 11th Annual Dayton, Ohio Chapter Symposium, 28 November, 1990.

[04] Raab, Frederick H., Ph.D., Algorithms for Position and Orientation Determination in Magnetic Helmet Mounted Sight System, US Air Force Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB OH, 1982, AAMRL-TR-82-045.

[05] Raab, Frederick H., Ph.D., and Brewster, C. C., Magnetic-Multipole Technique for Moveable-Scatterer Compensation, US Air Force Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB OH, 1988, AAMRL-TR-88-054.

[06] Sellers, John, Helmet Mounted Display Electronics for Evaluating Helmet Mounted Displays, Society for Information Display, International Digest of Technical Papers, Vol. 22, May, 1991

[07] Task, H.L., Kocian, D.F., and Brindle, J.H., Helmet Mounted Displays: Design Considerations, Advancement on Visualization Techniques, AGARDograph No. 255, October 1980.

THE EVALUATION OF PARTIAL BINOCULAR OVERLAP ON CAR MANEUVERABILITY: A PILOT STUDY

Brian H. Tsou, Ph.D.
Armstrong Laboratory
AL/CFHV
Wright Patterson AFB, OH

and

Beth M. Rogers-Adams
and
Charles D. Goodyear
Logicon Technical Services Inc
Dayton, OH

INTRODUCTION

An engineering approach to enlarge the helmet-mounted display (HMD) field-of-view (FOV) and maintain resolution and weight by partially overlapping the binocular FOV has received renewed interest among human factors scientists. Some evidence has been accumulated to suggest that any panoramic display, when binocular overlap is less than 100%, will be objectionable. As far back as 1962, overlapping the monoculars was used to obtain an ultrawide-field display with a 40-deg overlap for the Army.[1] Whether any visual problem was experienced with such a display was not known.

Panoramic HMDs, employing a similar approach, were later built as flight simulators.[2,3,4] It was reported[3] by CAE Electronics that luning or edge effect, described as two dark bands forming a distinct border in an otherwise uniform field around the central binocular overlap, was observed in these displays and attributed to binocular rivalry. McLean and Smith[5] reported that "partially overlapped fields (40-deg) are usually annoying, but with about 30 minutes of use most observers report not even being aware of the juncture areas." They also noted that head movements were increased with the partially overlapped HMD, during helicopter flights. Greene[6] reported a helicopter flight experiment in which pilots noticed increased illumination in the overlapped portion (20-deg) of the field. The pilots were slightly annoyed by the "luning" border, as well as by occasional minor eye fatigue. Greene noted that when higher distortion was present in a setup of 45-deg overlap, "airspeed/altitude performance decreased significantly, pilot ratings dropped significantly, and head motion increased." In addition, the pilots reported that "double vision, head aches and eye fatigue were common." All of the above studies used divergent optical arrangements (i.e., monoculars tilted outwards to create the partial overlap). Melzer and Moffitt[7,8] evaluated both divergent and convergent configurations (along with two other methods) for reducing edge effects. They found that there was less "luning" in the convergent display (i.e., with monoculars tilted inwards to create the partial overlap). Melzer and Moffitt stated that angular overlaps of at least 20-deg have been suggested but did not provide any reference.

Do edge effects and increased head motion affect performance? Melzer and Moffitt[8] reported that their studies show the ability to detect small targets is not affected by the edge. Kruk and Longridge[9] found no performance degradation in target detection, motion detection, or target tracking for a binocular overlap of 25-deg and 45-deg. There was degradation at the edges of the 25-deg overlap. Landau[10] found that a 17-deg overlap condition used in her recognition study produced degraded performance, while the 38-deg overlap did not reveal appreciable differences in accuracy or temporal performance. She also affirmed earlier reports of tendencies for head movement, binocular rivalry, and brightness variations. Whether conditions under which performance degradations were found simply reflect binocular probability summation[11], which is known to enhance binocular vision over monocular, will require further exploration.

It is evident, based on our brief literature review, that any panoramic display with a binocular overlap, less than a minimum amount, annoys the viewer, degrades performance, and elicits undesirable behavior. Whether these factors affect certain tasks performed in a dynamic environment is not clear and can not be adequately predicted. Our specific concern is the extraneous head motion that has been reported. These extraneous head motions, as suggested by the authors, were attributed to image distortion and alignment accuracy. From a system design point of view, it is important to establish whether increased head motion results in diminished performance and to verify if the cause for the increased head motion is display distortion[12]. If so, careful aberrational corrections of the HMD optics, as well as expensive and sophisticated image source distortion correction circuitry may be required. In this pilot study, the effect of varying distortion-free binocular overlap was evaluated with professional drivers maneuvering a car through an obstacle course. The experiment was conducted at the Transportation Research Center of Ohio (TRC), East Liberty, OH.

DIGITAL CLOCK
VCR
GRID COMPUTER
POWER INVERTER
GYRO PACKAGE
FIFTH WHEEL

CAMERA
RECEIVER
TRANSMITTER
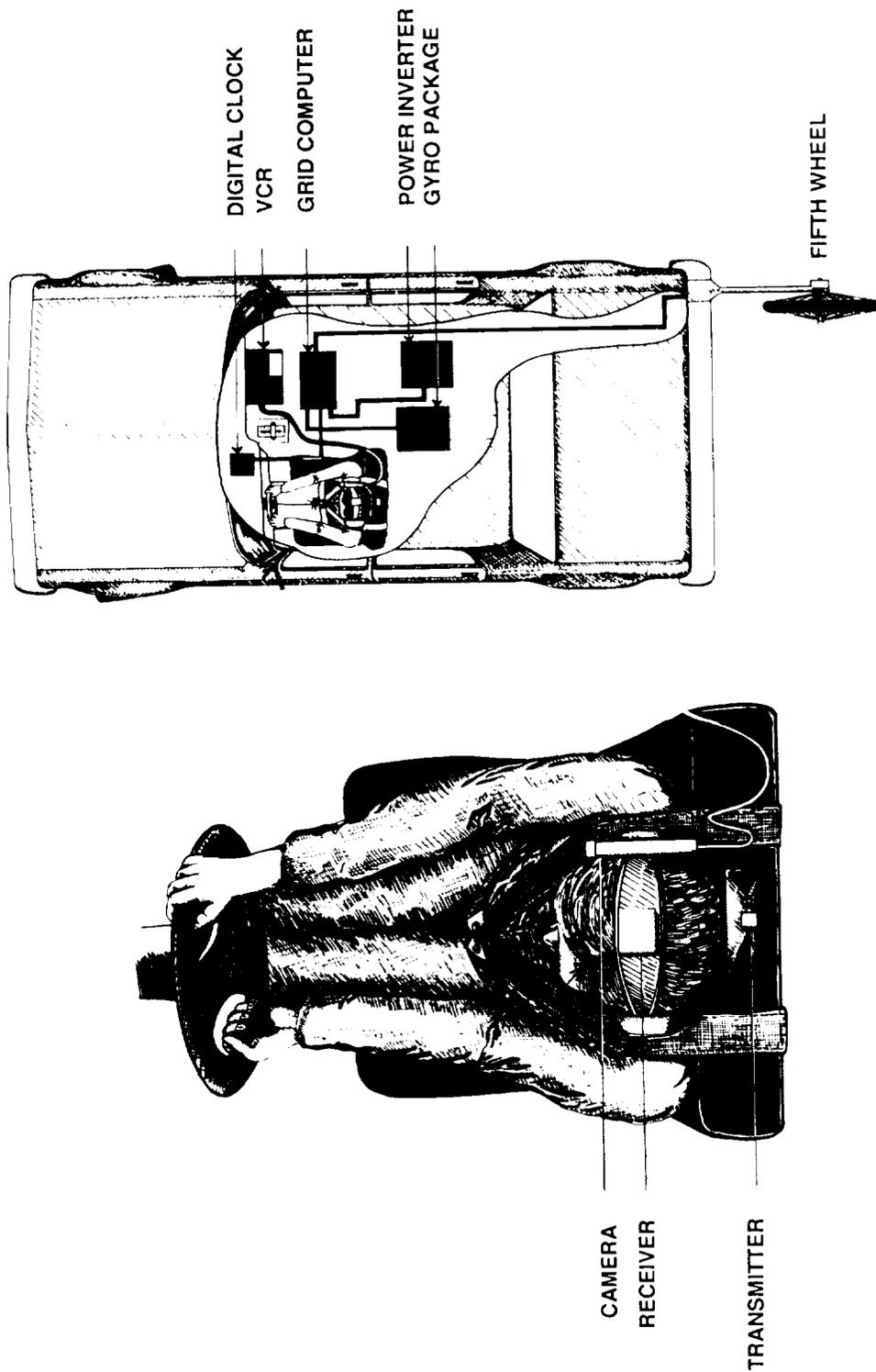
FIGURE 1.   EXPERIMENTAL APPARATUS FOR CAR AND SUBJECT

563

METHODS

SUBJECTS

Subjects were two male, age 34 and 39, highly experienced test drivers employed by the TRC. Both subjects were right handed and right eye dominate, with corrected 20/20 vision.

APPARATUS

The 1989 Chrysler Aries car, used for this experiment, was fitted with several data collection devices. Figure 1 shows the equipment layout inside the car. A Polhemus 3-Space tracker was used to record the subject's head movement in the azimuth and elevation planes. The head tracker's magnetic field transmitter was mounted over the driver's seat of the car. The receiver was mounted on a head band with ear cups and worn by the subject. Resulting from the fact that the on-board Grid computer could only accept analog data, an external digital-to-analog converter was constructed to translate the 3-Space tracker's digital head orientation angles into analog voltages in real time, before interfacing with the Grid computer (see Appendix 1). The Grid computer also recorded the vehicle dynamics, including car velocity and yaw rate, and provided the synchronizing timing signals. A calibrated fifth wheel was towed on the back of the car to measure car speed. The rate of car turn (yaw rate) was measured with a Humphrey gyro package. Car velocity, yaw rate, head azimuth and elevation angles were digitized by the Grid computer's peripheral data acquisition add-on board at 120 Hz (even though the head tracker was running at about 60 Hz). A Panasonic miniature color camera (CD1) was mounted on the side of the driver's head band pointing toward the front windshield of the car. The camera followed the driver's line-of-sight and recorded what the driver was viewing throughout each experimental trial. A microphone was also provided to record driver's comments, if any, as he maneuvered through the course. The outputs of the camera and microphone were recorded on a portable VHS VCR. The VCR recording was synchronized with the Grid computer at the beginning of each trial by pointing the camera and recording a digital clock display controlled by the Grid computer. The subjects were fastened into the car with two overlapping safety belts pulled tightly across their chests and waists to ensure little to no body movement, but free head movements.

PROCEDURE

Subjects performed a driving maneuverability task. The subjects were instructed to drive a car through an obstacle course as quickly and as accurately as possible. Some familiarization training was allowed on the obstacle course for both subjects. The time to complete each course was recorded using TRC's Alge stopwatch. An experimenter at the start gate signalled the start of each trial to the driver and to a second experimenter, by waving his/her arm. Simultaneously, the second experimenter, standing at the stop gate, started the stopwatch. The second experimenter stopped the stopwatch as the car passed the stop gate. The estimated accuracy was within a full second. A typical trial lasted approximately 30 seconds with a 5-10 minute interval between each trial. During this time interval, data were downloaded from RAM to the internal hard disk drive.

DESIGN

Independent measures included six varying obstacle courses and six different fields-of-view and overlaps (FOV/OVLP). The obstacle courses were located at one corner (approximately 600 x 800 feet) of the experimental area referred to as the Vehicle Dynamics Area (VDA). The test surface (asphalt) had a one-percent downward slope. The entire obstacle course was broken into three sections, marked out by pylons. The three sections became six courses by having the subjects drive in one direction, and then in the other. Figure 2 represents the layout of the courses.
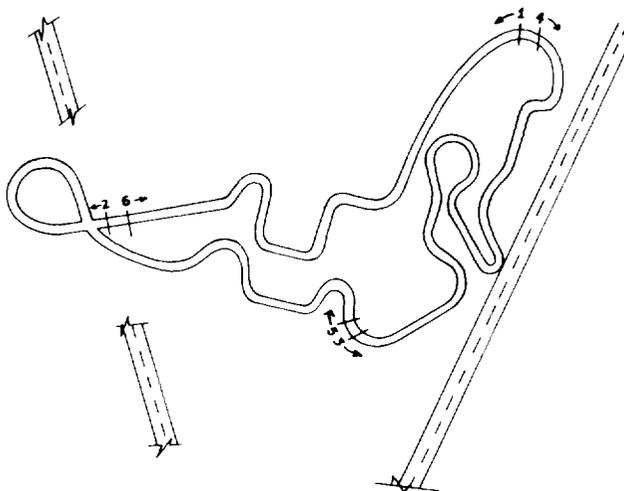

FIGURE 2. DIAGRAM OF THE SIX OBSTACLE COURSES

Both FOV and OVLP were simulated using baffles over a clear plastic eye-protective goggle, without any intervening optics, resulting in a *distortion-free, eye-limited* viewing condition. The individual subject's inter-pupillary distance (IPD) and eye relief (cornea to goggle) were measured to calculate the appropriate opening of a 4 x 3 aspect ratio rectangular format on the goggles for each FOV/OVLP. Black masking tape was used to cover the entire goggle except the desired opening in the front. The six FOV/OVLP conditions consisted of 180-deg (untapped goggle), 60-deg with 100% binocular overlap, 60-deg with 80% overlap (convergent and divergent) and 60-deg with 50% overlap (convergent and divergent). The definition used for the percent of overlap calculation is the amount of binocular overlap, divide by the total horizontal FOV (and not individual monocular FOV). This was to ensure a constant total horizontal FOV for all conditions, except the 180-deg which served as the baseline condition. Therefore, the 60-deg 80% overlap condition consisted of a binocular area of 48-deg, flanked by two 6-deg monocular areas. The 60-deg 50% overlap condition consisted of a binocular area of 30-deg, flanked by

two 15-deg monocular areas. Seventy-two trials were made up of a random ordering of two replications of six FOV/OVLP and six obstacle courses. The number of trials run each day was weather and time dependent. The whole VDA was covered with snow during the entire study, but the obstacle course was plowed and allowed to dry before experimental trials were run.

Dependent measures included course time (measured by stopwatches), error (displaced pylons), and head and vehicle dynamics (recorded with on-board equipment). The course time and error data were analyzed for the whole course, while car velocity, car turning rate, head azimuth velocity and movement, were extracted from the head and vehicle raw data only during the interval when the car was in and out of a turn. A program similar to a digital storage oscilloscope was developed on the Apple Macintosh, allowing the user to time-tag the head and vehicle traces in each trial for the beginning and end of each turn.

RESULTS

Analyses of variance were performed on the dependent measures, using subject, FOV/OVLP and course as the factors. The variability of the two replications was used as the error term. The Bonferroni procedure was used to make pairwise comparisons of FOV/OVLP, and course with an experimentwise error level of .05. There were no significant interactions between course and FOV/OVLP for all analyses.

Figure's 3 and 4 show course time and course errors as a function of FOV/OVLP. Course time analysis did not show any significant effect for FOV/OVLP (p=.907), but did show a significant effect (p=.0001) for course. Analyses on course errors indicated that FOV/OVLP and course had no significant effects, (p=.194) and (p=.076), respectively.

Significant main effects of FOV/OVLP were found for head velocity (p=.0001) and magnitude of head azimuth movement made during a turn (p=.0001). Paired t-tests showed a significant difference between the 180-deg and all other FOV/OVLP combinations. Figures 5-11 show head velocity, head azimuth movement, head turning time, car velocity, car turning rate, proportion of head directional change, and head leads car, respectively.

Significant main effects for course included head velocity, head azimuth movement, head turning time, car velocity, car turning rate (p=.0001) and the head leads car (p=.024). T-tests showed that in most cases, course six was significantly different from all other courses.
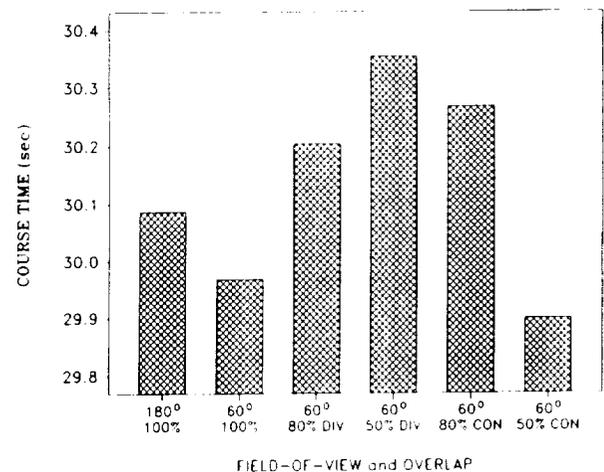

Fig. 3   Course Time plotted as a function of Field-of-View and Overlap.
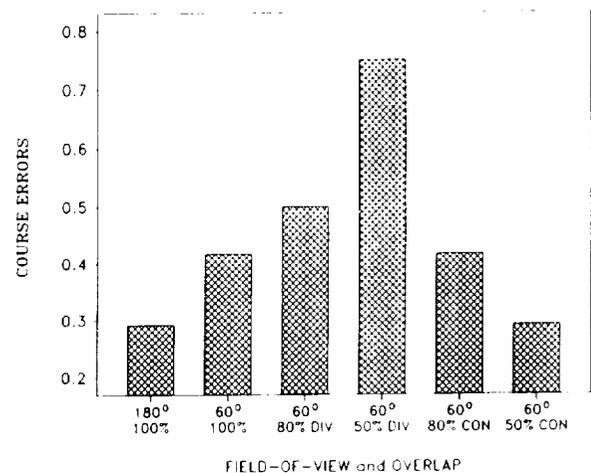

Fig. 4   Course Error plotted as a function of Field-of-View and Overlap.
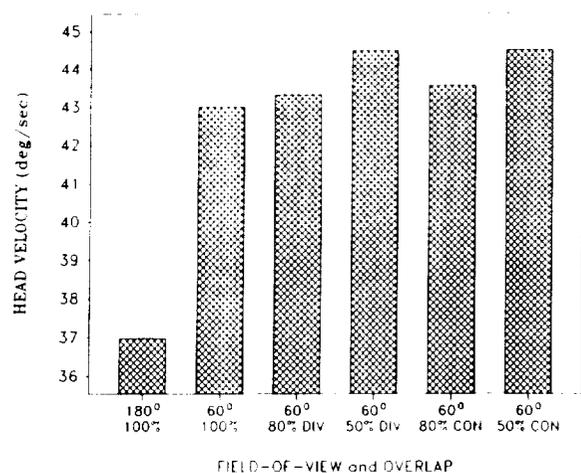
Fig. 5   Head Velocity plotted as a function of Field-of-View and Overlap.

565
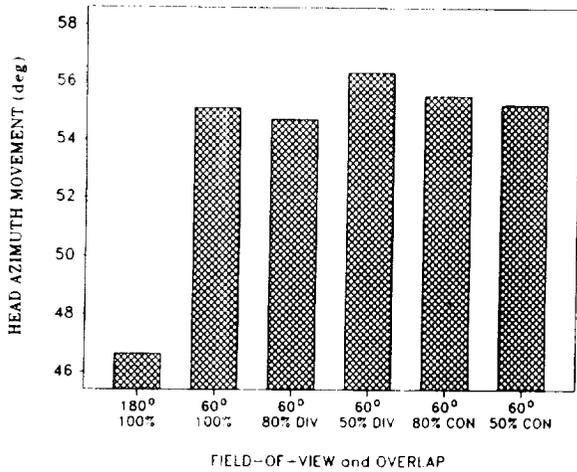
Fig. 6   Head Azimuth Movement plotted as a function of Field-of-View and Overlap.
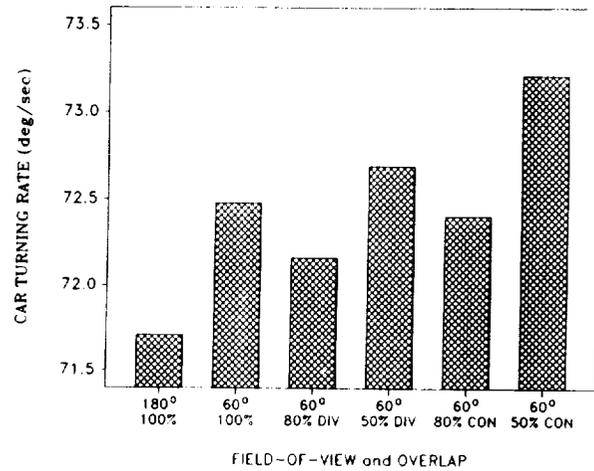

Fig. 9   Car Turning Rate plotted as a function of Field-of-View and Overlap.
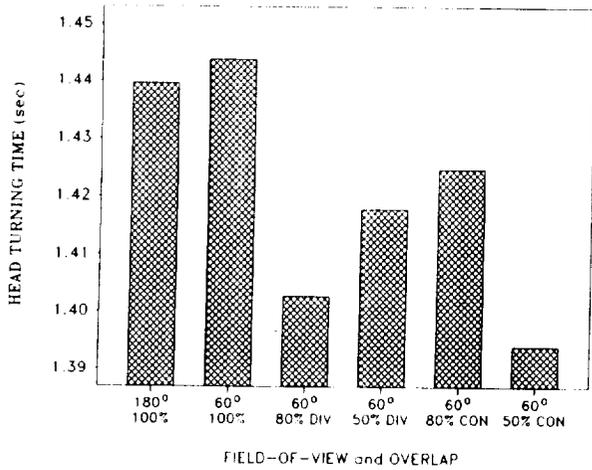

Fig. 7   Head Turning Time plotted as a function of Field-of-View and Overlap.
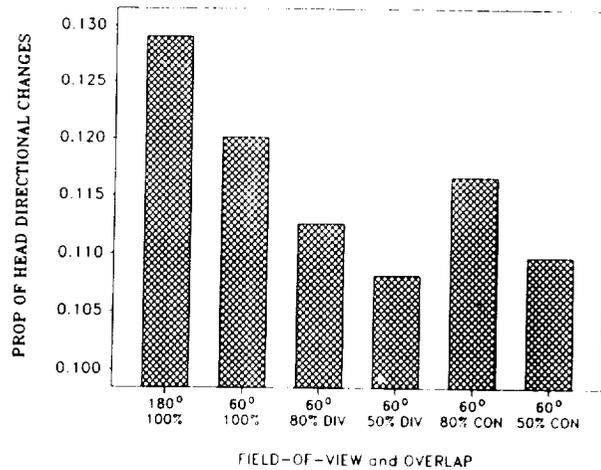
Fig. 10   Proportion of Head Directional Changes plotted as a function of Field-of-View and Overlap.
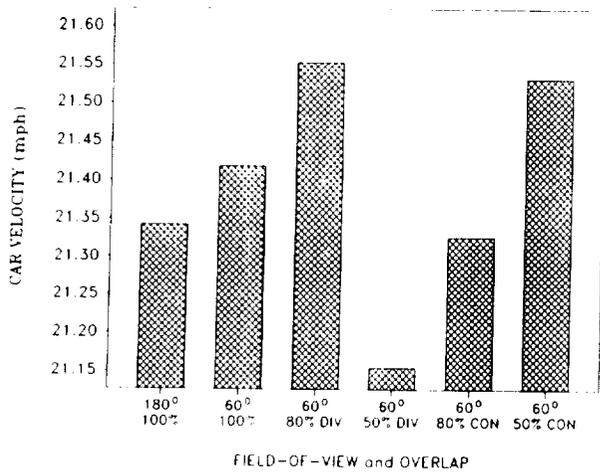
Fig. 8   Car Velocity plotted as a function of Field-of-View and Overlap.
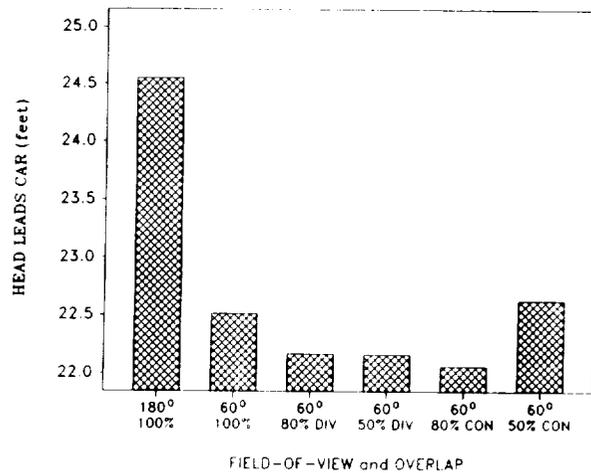
Fig. 11   Head leads Car plotted as a function of Field-of-View and Overlap.

566

CONCLUSIONS

Our subjects did not comment on any specifics regarding partial overlap conditions. It should be emphasized that because a rectangular format was used, the edge effect consisted of two "straight" borders. Subjects felt the 60-deg FOV did not inhibit them from performing the task in our study, but did not believe that 60-deg FOV would be sufficient for driving on busy city streets. Their performance on the course, as measured by course time and error, supported, at least the first half of their casual observations.

The major finding (Fig. 6) of any practical significance in this experiment is that across the 60-deg conditions, subjects moved their heads a greater distance (by about 5-degrees on each side) than in the 180-deg condition, presumably to compensate for the lack of FOV. Across all FOV/OVLP combinations, the elapsed times for completing a turn (Fig. 7), car velocity (Fig. 8) and turning rate (Fig. 9) were similar. Thus, this larger head movement translates directly into higher head velocity (Fig. 5). Though not significant, there is a slight trend suggesting that 50% overlap produced higher head velocities than 80%, which is higher than 100%. However, we can not rule out the association proposed by Greene[6]. Greene suggested that the higher head velocity is related to higher display distortion. A follow-on study, in which more subjects participate, and/or has more than one task, would be required to ascertain that partial overlap induces higher head velocities, even when there is no distortion. One indication that our subjects were not working as hard is that their head velocities were found to be in the 40-deg/sec range, compared to the 10-deg/sec range reported by Greene. One would assume, based on everyday experience, that a heavier workload would result in slower head motions.

Our head movement (head directional reversals) data (Fig. 10) did not support McLean and Smith's[5] observation. In fact, our data indicate that smaller overlaps produce less head movement. Again, because the differences were not significant, we can not say conclusively that decreased head motion is due to lack of binocular overlap.

Melzer and Moffitt[7,8] reported that there is less luning in convergent overlap[14]. We found no consistent differences between divergent and convergent overlap in terms of course time, error, head velocity or head movement. It is important to point out that in a convergent display, contrary to the divergent display and human binocular vision, the right eye will see more of the left (nasal) visual field and the left eye will see more of the right (nasal) visual field; subsequently if a target is moving from right to left, the left eye will detect the target before the right eye picks it up. This may cause confusion if the convergent panoramic display is not totally fused by the two eyes. We couldn't directly test this possibility in this study, but if we assume that the peripheral field is used in negotiating turns, then the measure of how soon the subject looks into the turn may detect that fine difference in the right/left eye reversal. Again, we found no significant difference (Fig. 11) between convergent and divergent configurations.

It is quite clear that our study, based on simple car maneuverability and two subjects, reveals differences in FOV, but nothing significant between binocular overlap levels and configurations. This tentatively indicates that some tradeoffs of binocular vision for a larger overall display FOV are acceptable. However, the need for further systematic experimentation in this area, to examine other relevant factors, is apparent.

APPENDIX

Append. 1[16]:
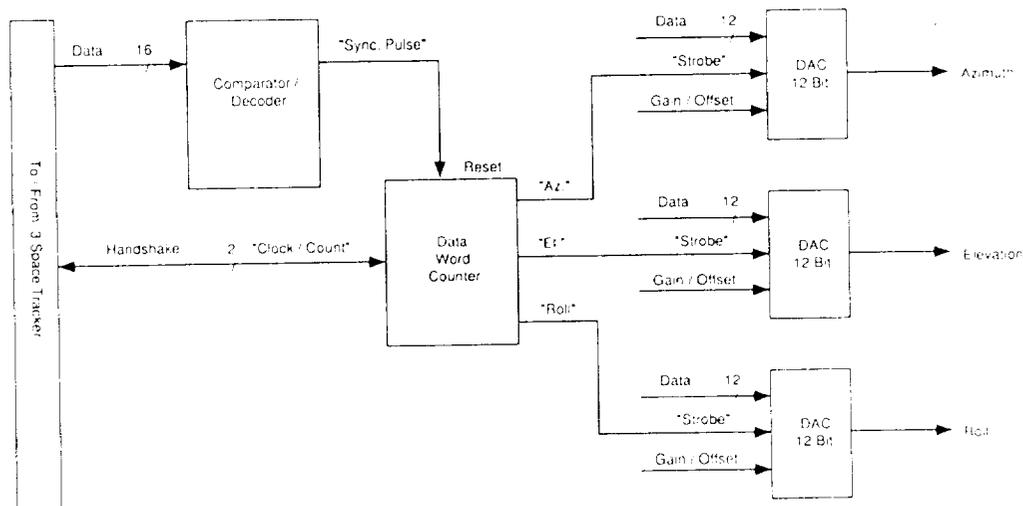
Block Diagram of the Digital-to-Analog Converter for 3-Space Tracker.

REFERENCES

1. Shenker, M., LaRussa, J., Yoder, P. and Scidmore W., "Overlapping-monoculars - An ultrawide-field viewing system," Appl. Opt. 1, 399-402 (1962).

2. Buchroeder, R., An optical analysis of the Farrand VCASS helmet-mounted display (AFAMRL-TR-83-072), Wright Patterson AFB, OH: Air Force Aerospace Medical Research Laboratory (1983).

3. CAE Electronics Ltd., Wide-field-of-view, helmet-mounted infinity display system development (AFHRL-TR-84-27), Williams AFB, AZ: Air Force Human Resources Laboratory (1984).

4. Welch, B., Kruk, R., Baribeau, J., Schlef, C., Shenker, M., Flight simulator: Wide-field-of-view, helmet-mounted, infinity display system (AFHRL-TR-85-59), Williams AFB, AZ: Air Force Human Resources Laboratory (1986).

5. McLean B. and Smith, S., "Developing a wide field-of-view HMD for simulators," SPIE Display System Optics, 778, 79-82 (1987).

6. Greene, D., Night vision pilotage system field-of-view (FOV)/resolution tradeoff study flight experiment report (NV-1-26), Ft. Belvoir, VA: Center for Night Vision and Electro-optics (1988).

7. Melzer, J. and Moffitt, K., "Partial binocular-overlap in helmet-mounted displays,", SPIE Display System Optics II, 1117, 56-62 (1989).

8. Melzer, J. and Moffitt, K., "An ecological approach to partial binocular overlap," SPIE Large Screen Projection, Avionics, and Helmet-mounted Displays, 1456, (1991).

9. Kruk, R. and Longridge, T., "Binocular overlap in a fiber optics helmet mounted display," The IMAGE III Conf. Proc., 363, 363-377, Air Force Human Resources Laboratory (1984).

10. Landau, F., "The effect on visual recognition performance of misregistration and overlap for a biocular helmet mounted display," SPIE Helmet-mounted Displays II, 1290, 173-184 (1990).

11. Arditi, A. (1986). Binocular Vision. In K.R. Boff, L. Kaufman, and J.P. Thomas (Eds.), Handbook of perception and human performance (Vol. 1, Chap. 23). New York: John Wiley and Sons.

12. Because most HMDs use F Theta mapping, there will always be a requirement for distortion correction. The system designer needs to balance between minimizing any residual distortion to a reasonable level and the costs. Another point relating to the distortion that should be made is that tilting of the monoculars to create the panoramic display does not result in a trapezoidal distortion, as concluded by at least two studies[2,7]. The reason for this is that most HMDs will be utilized as a one power telescope[13].

13. Shenker, M., "The power of one power," SPIE International Lens Design, 1354, 647-653 (1990).

14. The object/background depth relationship between convergent and divergent overlaps brought forth by Melzer and Moffitt[7,8] as the explanation of less luning in convergent display is rather intriguing and worth further exploration. However, there is an alternative hypothesis: In the convergent display both edges lie in the nasal visual field, therefore, falling on the temporal retina which is not as well tuned to binocular disparity[15] as compared to the nasal retina. Perhaps, they are less sensitive to binocular rivalry as well, which may account for the less "luning" in a convergent display. An experiment is planned to examine this hypothesis in the future.

15 Pettigrew, J., "The neurophysiology of binocular vision," Sci. Am. 227, 84-95 (Aug 1972).

16. Detailed schematics will be available upon request to Brian Tsou, AL/CFHV, WPAFB, OH 45433.

I

Three dimensional tracking with misalignment between display and control axes

Stephen R. Ellis
NASA Ames Research Center
Moffett Field, CA. 94035

Mitchell Tyler*, Won S. Kim** and Lawrence Stark
School of Optometry and School of Engineering
University of California, Berkeley
Berkeley, California 94720

Abstract — Human operators confronted with mis-aligned display and control frames of reference performed three-dimensional, pursuit tracking in virtual environment and virtual space simulations. Analysis of the components of the tracking errors in the perespective displays presenting virtual space showed that components of the error due to visual-motor misalignment may be linearly separated from those associated with the mismatch between display and control coordinate systems. Tracking performance improved with several hours practice despite previous reports that such improvement did not take place.

I. INTRODUCTION

The design task for telemanipulation displays is complicated by the fact that the three-dimensional data they depict may entail viewing from nonoptimal directions. The viewing projection is described by a number of parameters each of which can substantially alter the appearance of the resulting image. Since the purpose of the display is the accurate depiction of spatial information and successful manipulation of objects, designers must first understand the effects of these display parameters on operator perception and control in the work space. Studies of the effects of some of the major projection parameters, i.e. viewing direction and field of view angle (FOV), on the perception of exocentric direction have been reported recently [1] [2] [3] [4]. Other studies of manipulation and control of telemanipulated objects also have been reported over the past 20 years [5] [6] [7] [8][9] [10] [11].

The present experiments and analysis extend these investigations with studies of pursuit tracking of virtual objects moving irregularly in three dimensions and focus on the spatial components of the tracking error. In these experiments the FOV angle is defined as

the visual angle subtended by the viewport as seen from the geometric center of the projection. The azimuth and elevation of the viewing vector are defined by angles between the world coordinate system and an extension of the viewing vector to a reference point. In general, the FOV was correctly matched by placing the subject's eye at the correct station point while misalignments between display and control axes were introduced by rotating the azimuth of the viewing direction. These misalignments were used to produce a decrement in human tracking performance in a manner similar to that of studies by Bernotat [8]. The resulting pattern of degraded performance, which has been found to be a function of the amount of misalignment, provides a basis for examining Bernotat's claim that subjects cannot adapt to the misalignment. In addition, errors in tracking will provide data to assess the suitability of integrated measurement of tracking error versus analysis of the components of the error.

II. EXPERIMENT 1

A. Three-dimensional Tracking with Perspective Display and Joystick Controllers

B. Methods

This experimental setup was identical to that described in an earlier paper [9]. The basic task was to control two, 2-axis joysticks to track the projection of a small diamond-shaped target that moved irregularly by sums of sines in three dimensions. The forcing functions in all three dimensions were determined so their spectrum resembled that of a first order, low pass filter with a cut-off at 0.1 Hz. Five different forcing functions were precomputed and randomly selected for a particular two minute tracking run. Subject's were given a ten second warm-up at the outset of each two-minute tracking run. The sampling and display frequency was 40.96 Hz.

The perspective projection on the screen was computed by an 11/23 computer and displayed on a HP1345A stroke monitor. The cursor was a small cross displayed at half brightness. It was controlled in horizontal position by two axes of one joy-stick and in vertical position by the fore-aft axis of the other joystick. The tracking environment, a 5x5 line grid, was oriented to appear parallel to the floor and provide a spatial reference for pursuit tracking. Four tracking errors were computed: Integrated normalized root mean square (RMS) error between the target and cursor, and the separate x, y, and z components of this error. In normalized RMS tracking error a value of 1.0 corresponds to the error expected if no tracking is attempted while the cursor remains centered. If t is the target position and c is the cursor position and $x = (x,y,z)$, the computed errors were:

$$RMS = \frac{\sqrt{\sum_i^N |x_c - x_t|^2}}{N}$$

$$RMS_x = \frac{\sqrt{\sum_i^N (x_c - x_t)^2}}{N}$$

Similarly for RMS_y and RMS_z
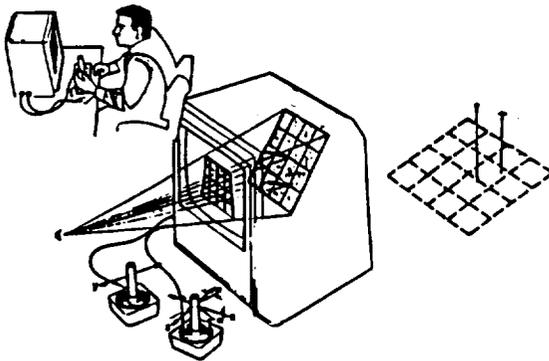
Figure 1. Schematic setup for 3D tracking using two 2 axis joysticks with a target on a perspective display.The control axes of the right joystick could be rotated into alignment with the viewing vector. To the right of the schematic is a tracking display image with a 45° rotation of azimuth and a pitch down of 45°.

The view vector was pitched down 45° and the azimuth was rotated through 9 positions either cw or ccw to provide a variable amount of display and control misalignment. The central axis of the grid represented the X and Z coordinates of a Cartesian coordinate system. The experiment was conducted in a darkened room but the contours of the display case and other laboratory equipment were still easily visible and well into photopic vision. Subject's were seated 40 cm in front of the display screen at the correct geometric eye point for the projection as shown in Figure 1.

The principal manipulation of the experiment was the orientation of the control axes of the joystick and the orientation of the viewing vector. In one condition it remained constant (unrotated condition) as the viewing vector was rotated, thus allowing a display-control misalignment to develop. In the second condition (rotated condition) the control axes were rotated so as to maintain alignment with the display axes as the view vector was rotated (See Figures 1 and 3). The comparison of tracking performance in these two different conditions would allow the investigators to isolate of that tracking component due to the purely visual consequences of rotating the viewing vector, shown in the rotated condition, from the consequences of visual-motor misalignment. The purely visual-motor component would be seen as the difference between the unrotated and the rotated conditions. The direction of rotation (cw/ccw) was reversed for alternate subjects. Half the group experience the rotated joystick axis condition first and half the unrotated condition first. The specific sequence of rotations presented was randomized before each experimental condition was run with a particular subject.

C. Subjects

Six male subjects enroled in a mechanical engineering course at U.C. Berkeley served as subjects. They had all had sufficient experience with the tracking task to have reached asymptotic performance with joy-sticks and the general task, but they had had no practice with all the misalignment conditions.

D. Results

Inspection of the data averaged across subjects in Figure 2 for the unrotated, normal conditions shows two significant features. First, the best tracking appears to be at an azimuth of 0°. There appears to be very little effect on tracking performance until the azimuth angle exceeds roughly 50°. Thereafter, the error increases to a maximum at about 125° thereafter decreasing to 180°. Rotation of the control axes to maintain visual-motor alignment had appeared to eliminate totally the tracking disturbance. However, further analysis of the components (Figure 4) of the tracking revealed a residual tracking error, attributable purely to the visual aspects of the track-

Integrated RMS Tracking Error
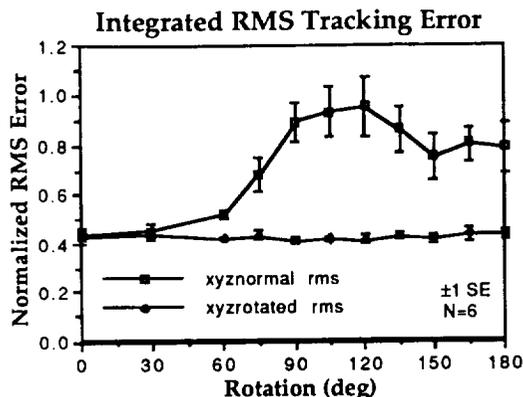
Figure 2. Comparison of RMS tracking error for rotated and unrotated tracking conditions.

ing. The visual component of the tracking along an axis is most accurate if the axis is viewed from an orthogonal direction. The variable component of tracking is roughly proportional to the sine of the angular deviation from the optimal view. Thus, the peaks in the x and z components tracking error were 90° apart. Significantly, if the visual component revealed by the rotated condition is subtracted from the unrotated condition, the x and z components of the tracking are now almost identical. This similarity suggests that the visual component and the visual-motor component of the tracking errors may be linearly separable.

III. EXPERIMENT 2

A. Three-Dimensional Tracking in Virtual Environment with Anthropomorphic 3-dof controller

In contrast to Bernotat's report [8] that the performance disruptions caused by display-control misalignments do not dissipate with practice, we have observed in a previous three dimensional tracking experiment [2] evidence that tracking does improve with several hours practice. To verify that this training effect also takes place in other tracking environments, a tracking experiment similar to Experiment 1 was conducted in a virtual environment using a hand position sensor. In this case the subject used a direct mapping of hand position to cursor position and did not have to learn the motor mappings involved in use of a joystick to control a cursor. The more natural mode of cursor control provided by a hand tracker could be expected to facilitate adaptation to the display-control misalignment.

B. Methods

The earlier tracking studies [2] as well as Experiment 1 used 5 distinct target disturbance functions with the identical frequency content. The target functions were randomly selected for each run without the subject's knowledge. This second

experiment which copied the dynamic and visualcharacteristics of the previous experiments, used unique forcing functions for each run. Use of unique functions for each run can rule out the remote possibility that subjects can learn the specific paths in the fixed target function.

In the present study the tracking environment was presented as a virtual stereoscopic image via a head-mounted, CCD (charge-coupled device) display developed for technology demonstration purposes at the Aerospace Human Factors Division at the Ames Research Center [12]. For the tracking experiment the overall sampling and display rate was approximately 30 Hz. Other experimental conditions were comparable to those in Experiment 1. The subject's head position was tracked electromagnetically during the experiment and used to approximately inertially stabilize the visual display. They were not, however, encouraged to walk about after first selecting a viewpoint that appeared subjectively optimal. Since the entire grid display was visible without head movement, none was required during the tracking. But both subjects made small (<20°) rotational movements anyhow. The subjects stood directly below a Polhemus electromagnetic 6 DOF tracker which was mounted on a plastic support 7 ft above the ground. Based on a calibration test, the reach volume of both subject's right hands was within a volume in which the tracker performed close to factory specifications. The tracking space was scaled 1:1 within the synthetic environment and the gain for displacement in all axes was set at 1.0. The subjects tracked targets under the nine different visual-motor misalignment conditions with two replications in each set for a total of 18 randomly ordered tracking runs per group. Each subject was given 3 groups for a total of 54 runs. Subjects were given rest period of approximately 10 minutes between groups. Before data collection began subjects were given several practice runs and informal training under unrotated tracking conditions to be sure they understood the tracking task and that the head-mounted display was properly positioned. As in Experiment 1, the first 10 seconds of tracking was considered a warm-up period and not analyzed

C. Subjects

Two male subjects, ages 18 and 63 years, who were laboratory personnel at Ames participated in the experiment.

D. Results

The individual tracking RMS errors from both subjects (Figure 6) indicate a similar performance patterns compared to earlier data. As the misalignment between display and control axes is increased, tracking generally deteriorated. The relative amount of the deterioration seemed to be less and maximum degradation around 125° of misalignment is
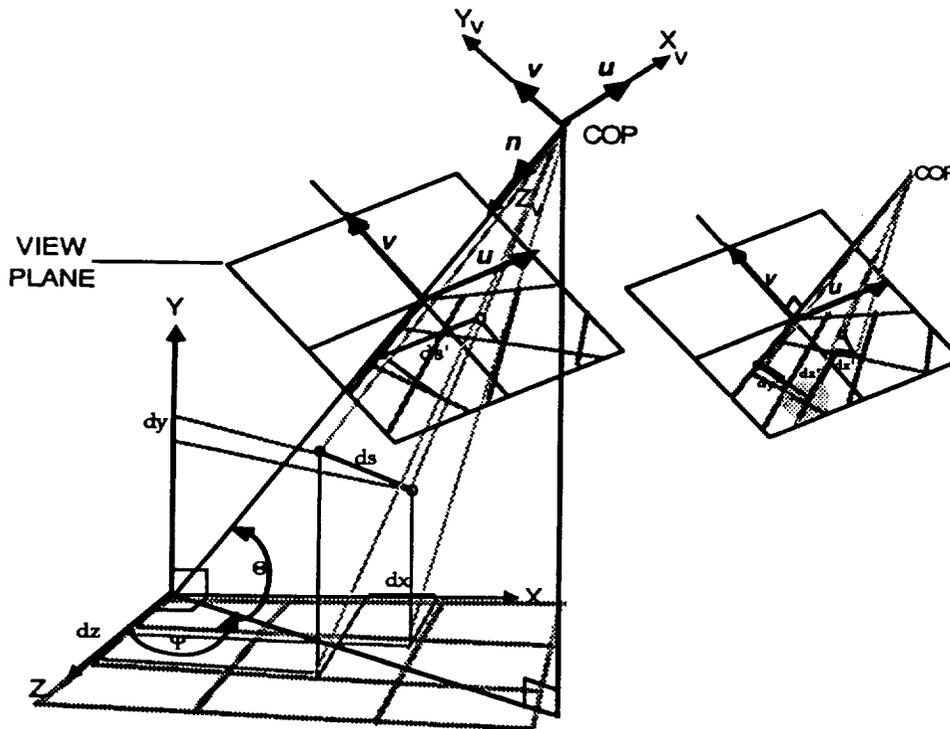
Figure 3. A target at ● is disturbed in 3 dimension and tracked by cursor at O . The viewing is with azimuth of Ψ and elevation of Θ from COP. Integrated instantaneous tracking error and its x,y,z components are ds, dx,dy, and dz respectively. Projections of these errors are on the viewplane, as ds', dx', dy' and dz' The view plane has been redrawn to the right for clarity. The disk of dots on the right view plane represents line of sight measurement noise.
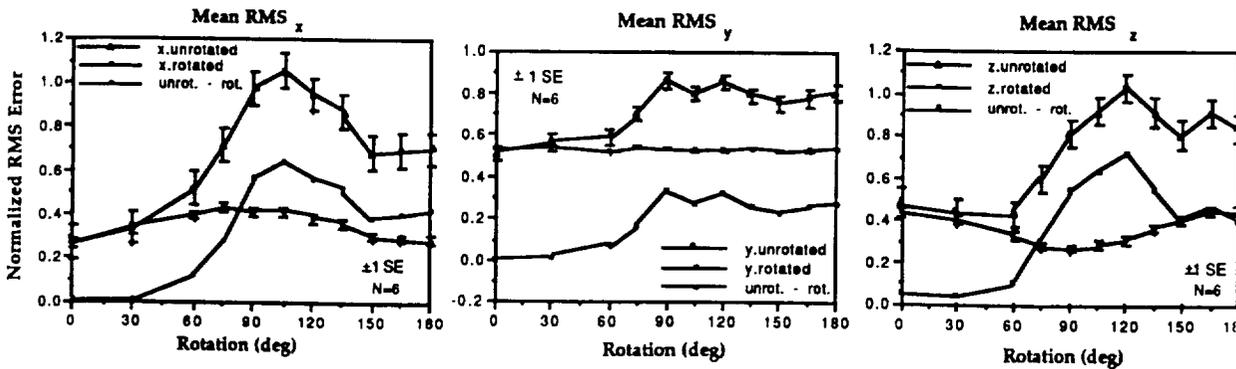

Figure 4. Across subjects means (n=6) of integrated RMS and its x,y,z components for unrotated (upper traces) and rotated control-axes (middle traces) during 3D tracking on perspective displays. The lower trace in each figure results from subtracting the rotated conditon from corresponding tracking results in the unrotated condition

not as pronounced as in earlier results. Tracking performance of the two subjects does not appear to stabilize until the third group of tracking runs. Differences in the absolute RMS values reflect differences in individual tracking capabilities as well as the poor display resolution and temporal response of the head-mounted display. Significantly, despite many technical and procedural differences with earlier studies, both subjects confirm tracking performance improvement. In fact, the peak disturbance around 125° of misalignment almost vanishes in both subjects (Also see Figure 7 from an earlier experiment)

E. Discussion

Difference in dynamics used to control the cursor as compared to Bernotat's earlier work may account for the above confirmation that subjects can adapt to the misalignment of the control frame of reference . Bernotat simulated a 2nd order system which gave the cursor a significant sluggishness causing a delay between the subjects actions and his observation of an effect. This delay could disturb subject's ability to learn a compensatory motor strategy that would improve from session to session. Another significant difference was that Bernotat used compensatory tracking and we used a pursuit paradigm. The compensatory tracking may have obscured the effects of the subject's control inputs by mixing them with the targets own motion. This mixing increases the difficulty he would have in identifying the effects of his control inputs on the target.

Figure 5. Subject wearing the virtual environment display while performing 3D tracking.

A common feature of the several experiments we have conducted on display-control misalignment has been that the worst tracking performance seems to be at at misalignment angle somewhat greater that 90°. This finding is even suggested by Experiment 2 in that the peak error seems to be at rotations greater than 90°, perhaps around 125°. The

peak may remain evident after considerable practice Experiment 1 again confirms this point but more subject's are needed to confirm it in Experiment 2.
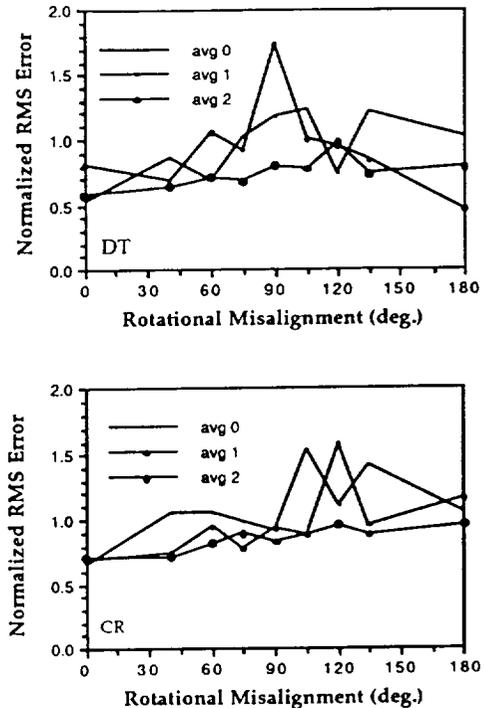
Figure 6 Two Individual subjects integrated RMS 3D tracking data averaged within each group of 18 runs.

The observation that the poorest tracking performance may be at azimuth rotations near 125° invites speculation. The subjects ability to compensate for the viewing rotation of the reference grid with respect to the control axes of the joystick may involve a mental rotation of the grid. Cooper and Shepard [13] have shown that the time required for such mental rotation varies in proportion to the size of the angle of rotation. If this effect occurs in manual tracking, a phase lag between the target trajectory and the human response should be detected that is proportional to the azimuth angle used to generate each display condition.

But why should the maximum tracking error occur at a rotation greater than 90°? At 90° the subject could rotate either clockwise or counterclockwise to align with the control axes. Depending upon how the subject visualizes his mental rotation, one or the other of these directions of rotation will also require a control reversal. Clearly, the subjects will choose the rotation that does not require a reversal. But what should they do at 91°? Assuming that the control reversal will add additional time lags and further degrade

573
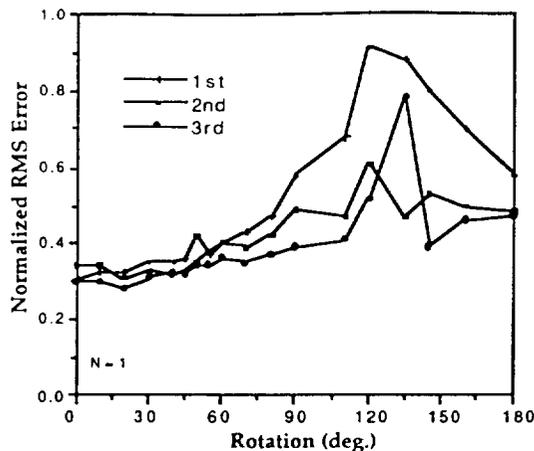
Figure 7. The results of multiple two minute tracking runs from a single well-practiced subject who experience display-control rotational misalignments from 0 to 180° in an earlier experiment using a perspective display [2]. The data show both evidence of learning across groups of runs (1,2,3) and a peak disturbance in tracking near 125°.

tracking performance, one may speculate that the subject will choose to make mental rotations greater than 90° to avoid having to introduce the reversal.

At some angle, however, a point may be reached at which the delay due to a control reversal may be acceptable because the alternative rotation becomes sufficiently small. Thereafter, the total delay should be a decreasing function of azimuth; as the azimuth increases, the amount of required mental rotation decreases. Another consequence of this analysis is that the intra and intersubject variability ought to be at a maximum near the point of maximum RMS error since this would be where most equivocation regarding the direction of corrective rotation should occur.

Understanding of the consequences and causes of spatial interpretation and control of three-dimensional cursors will assist in the design of three-dimensional spatial instruments for use in aerospace applications i.e. [14].

IV. CONCLUSIONS

1) Manual three dimensional pursuit tracking errors produced by display-control rotational misalignments have two linearly separable components: a purely visual component and a visual-motor component. Both components may be independent influences on tracking performance.

2) Human subjects can simultaneously adapt to a variety of display-control misalignments if using position control during pursuit tracking with a

simulation update rate of at least 30 Hz. This capability will enable trained operators to quickly adapt to changes in the position and orientation of viewing cameras during teleoperation and telemanipulation.

ACKNOWLEDGEMENTS

REFERENCES

[1] McGreevy, Michael W. and Ellis, Stephen R. (1986) Th effect of perspective geometry on judged direction in spa tial information instruments. *Human Factors*, 28, 439-456.

[2] Ellis, S.R., Kim, Won Soo, Tyler, Mitchell, McGreevy, M W., Stark, L. (Nov., 1985). Visual enhancements for perspective displays: perspective parameters. *Proc. of the Inter. Conf. on Syst. Man and Cyber..* IEEE Catalog # 85CH2253-3, 815-818.

[3] Tharp, Gregory and Ellis, Stephen R. (1990) Modelling Changes in Errors of Perceived Direction in Perspective Displays *NASA TM 102792.*

[4] Barfield, Woodrow, Lim, Ravael, and Rosenberg, Craig (1990) Visual enhancement and geometric field of view as factors in the design of a 3-dimensional perspective display. Proceedings of the Human Factors Society, 1470-1473.

[5] Smith, Karl U. and Smith W. M. (1962) *Perception and motion: an analysis of space-structured behavior.* Philadelphia, W.B. Saunders.

[6] Smith, Thomas J. and Smith, Karl U. (1987) Feedback control mechanisms of human behavior. In *Handbook of Human Factors*, G. Salvendy (ed.) New York, Wiley.

[7] Vertut, Jean and Coiffet, Philippe (1986) Robot Technology: teleoperations and robotics: evolution and development Vol 3, Prentice Hall, Englewood Cliff, NJ.

[8] Bernotat, Rainer K. (1970) Rotation of visual reference systems and its influence on control quality. *IEEE Trans on MMS*, MMS-11, 129-131.

[9] Kim, W. S., Ellis, Stephen R., Hannaford, Blake, Stark, L. A quantitative evaluation of perspective and stereoscopic displays in three axis manual tracking tasks. (1987) *IEEE Trans. on Sys. Man and Cyber.s, SMC-17*, 61-71

[10] Hannaford, Blake (1990) Kinesthetic feedback techniques in teleoperated systems in *Advances in control and dynamic systems,* C.T . Leondes (ed.)Academic Press, NY.

[11] Cunningham, H. A. (1989). Aiming error under transformed spatial mappings reveals spatial structure of visual-motor maps. *Journal of Exper. Psych.: Human Percep. and Perf.,* 15, 493-506.

[12] Fisher, Scott S., McGreevy, Michael W., Humphries, J., Robinett, W. (1986) Virtual environment display system. ACM 1986 Workshop on 3D Computer Graphics, Chapel Hill , NC October 23-24, 1986.

[13] Cooper, L.A., Shepard, R.N. (1978) Transformations on representations of objects in space. In E.C. Carterett and M.P. Friedman (Eds.) Handbook of Perception, Vol III: Space and Object Perception. New York, Wiley .

[14] Ellis, Stephen R., McGreevy, Michael W., Hitchcock, Robert (1987) Perspective traffic display format and airline pilot traffic avoidance. *Human Factors, 29,* 371-382.

ANGULAR RELATION OF AXES IN PERCEPTUAL SPACE

Urs Bucher

Department of Psychology, Biomathematical Section, University of Zurich, Switzerland
and
NASA Ames Research Center, Human Factors Research Division,
Mail Stop 262-2, Moffett Field, California 94035-1000, USA

ABSTRACT

The geometry of perceptual space needs to be known to model spatial orientation constancy or to create virtual environments. To examine one main aspect of this geometry we measured the angular relation between the three spatial axes.
We performed experiments consisting of a perceptual task in which subjects were asked to set independently their apparent vertical and horizontal plane. The visual background provided no other stimuli to serve as optical direction cues. The task was performed in a number of different body-tilt positions with pitches and rolls varied in steps of 30°.
The results clearly show the distortion of orthogonality of the perceptual space for non-upright body positions. Large interindividual differences were found. Deviations from orthogonality up to 25° were detected in the pitch as well as in the roll direction.
Implications of this non-orthogonality on further investigations of spatial perception and on the construction of virtual environments for human interaction will also be discussed.

INTRODUCTION

Space constancy, achieved by space transformations continually performed in the CNS, is an amazingly reliable ability providing appropriate interactions with the environment. Three different sources of information are used to determine the transforming operation: 1) visual direction cues, 2) somaesthetical direction cues and 3) vestibular direction cues. To model spatial orientation constancy and to create a virtual environment, we have to analyze each of these cues separately and learn about their interaction. Doing this, we find out that it is not always as accurate as one might expect. In the present study, we tried to perform experiments in which visual direction cues were eliminated, to vary mainly the vestibular stimulation while reducing somaesthetical direction cues as much as possible. While there is abundant data on the perception of the vertical, there are relatively little data on the whole perceptual space and the angular relationships of its axes (Bischof, 1974; Bucher, 1988). The reason for this lack of data might be the assumption that the internal representation of space is orthogonal and, therefore, that measuring the perceived vertical also provides data for the perceived horizontal.
This paper provides evidence to suggest that this assumption may be invalid.

METHODS

Apparatus.

Our apparatus allowed us to tilt human subjects into every desired body position respective to gravity (see fig. 1). The cockpit (diameter 110 cm; width 62 cm) in which the subjects were placed could be turned forward and backward in order to vary the pitch dimension. By turning the whole frame in which the cockpit is suspended, we were able to tilt the subject sideways thus varying the roll dimension. Both possible movements could be performed independently as well as in combination. The actual position of the cockpit and the frame was measured electronically with a accuracy of 0.1°.

To reduce extra-otolith postural influence on space perception the subject was placed in a seat of inflatable pillows. This ensured that the subject remained in a fixed position and

afforded a more constant and equal distribution of the pressure that he/she experienced. Stabilized by an easily removable bite-board, the subject looked through binoculars. To ask subjects about their perceived vertical/horizontal an adjustable luminous line/ring was presented. By using a UV light-source and a black background the stimulus seemed to be free floating in space. The device had two degrees of freedom, which could be manipulated by the subject with two control knobs mediating the two step motors. An onboard camera, equipped with a macro optical lens and connected to a video system, was used to independently monitor each eye in order to determine the ocular counter-rolling (procedure described detailed in Bucher, Heitger, Mast & Bischof (1990)).

The entire apparatus was remote controlled by a PDP 11/73. Each experimental session could be prepared off-line for a subsequent fully automatized performance of the experiment.
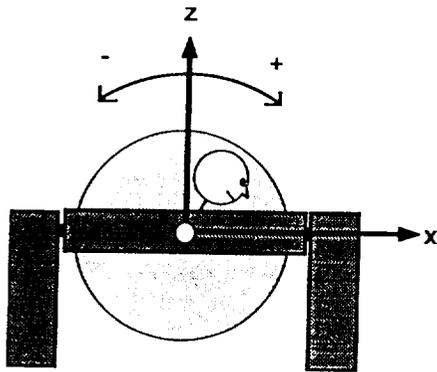

Figure 1a : Pitch
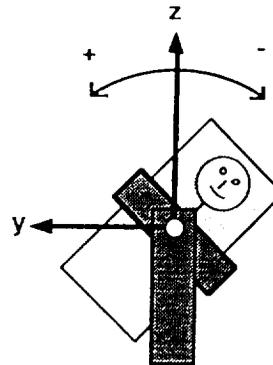
Figure 1b : Roll

Figure 1
Apparatus used to stabilze subjects at various pitch and roll body tilts.
It shows the orientation of the pitch dimension (turning the cockpit, fig. 1a)
and the roll dimension (turning the frame, fig. 1b)

Experimental setting.

During each session the subjects were tilted in total darkness to 7 different consecutive body positions from 0° down to 180° in steps of 30°; 2 sessions for roll variation (right or left ear down) and 2 sessions for pitch variation (tilting forward or backward).

In every body position they had to perform the following set of tasks: a) place the luminous line according to the apparent vertical (the line was randomly preset in darkness with a deviation of about 20° in pitch and roll from the objective vertical), b) verify this initial placement twice and, if necessary, adjust this line position (after disappearing and

576

I

reappearing), c) repeat steps a) and b) with a luminous ring to place according to a horizontal plane. This set of tasks was performed 3 times.

Each session took 55 to 75 minutes from boarding the cockpit. (For a more detailed description see Bucher (1988)). Four subjects took part in the experiments: two females and two males , between 25 and 40 years. Their state of health was checked by standard medical testing.

RESULTS

Roll condition

Due to large *inter*individual differences, as they can be observed often in perceptual experiments, the results will be presented for each subject separately.

Figure 2 displays roll deviations of these settings for the two roll conditions (+180 right ear down, -180° left ear down). For every body position two means were calculated: One for the apparent vertical, which was given by the settings of the luminous line,

and one for the apparent horizontal, for which the normal on the plane, described by the settings of the luminous ring, was taken. The solid line represents the values of the vertical, the dashed of the horizontal. The functional characteristics of both curves are about the same whereas they differ clearly in amplitude. More striking, this fact is demonstrated in figure 3, which shows the absolute angular differences between the apparent vertical and the normal on the horizontal in the same conditions (solid line). If the perceptual space strictly would underlie the concept of orthogonality this angular difference would be zero and consequently the solid line identical with the x-axis. Since the standard deviation increases considerably at body tilts larger than 90° it was included in the graph as a reference curve (dotted line); it represents the double standard deviation as a statistical criteria. In general it can be shown that the right angle between the apparent horizontal plane and the apparent vertical is maintained no longer as soon the body is tilted away from its upright position. Although mostly below 10°, deviations as large as 25° are found in body tilts over 90°.
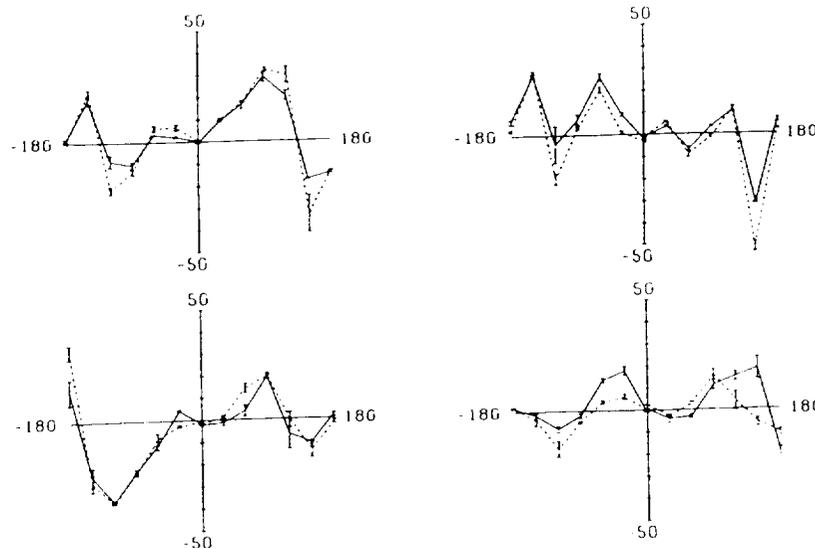
Figure 2
Deviations in the roll dimension under roll conditions:
The settings of the apparent vertical (solid line) and the normal on the apparent horizontal (dotted line).Each crosshair presents data of two independet experiments: roll right ear down (0° to 180°) and roll left ear down (0° to -180°).

577
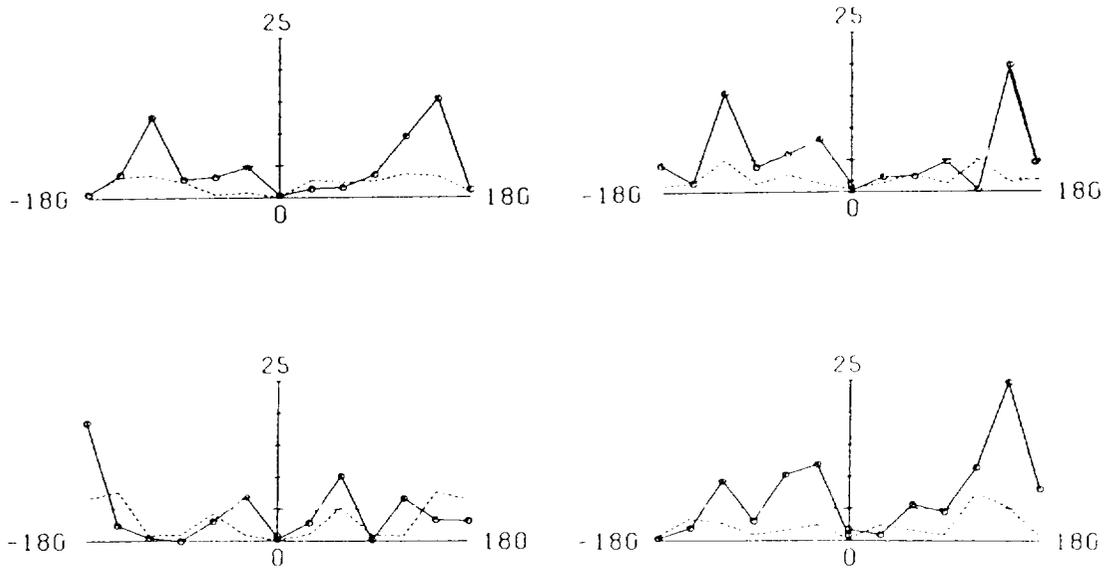
Figure 3
Deviations from the right angle between the apparent vertical and horizontal
in the roll dimension under roll conditions from 0° to 180° (right ear down) and
from 0° to -180° (left ear down). Subjects 1 to 4.
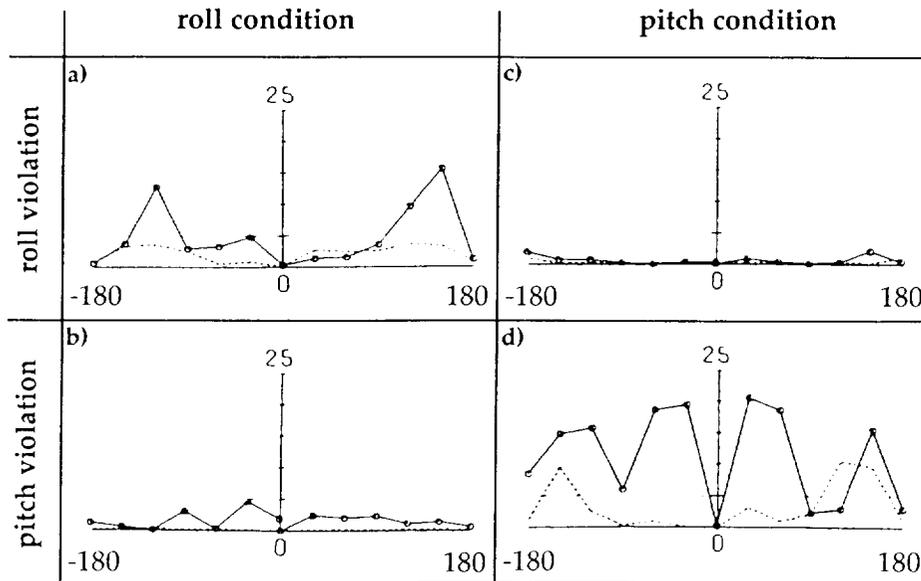
Figures 4a-d
Subject 1 in pitch and roll conditions:
Pitch and roll violations of orthogonality

578

Pitch condition

When tilting the body in pure pitch direction similar general characteristics can be observed. The spatial distortions can be broken up in two relevant violations of orthogonality: roll-violation component and a pitch-violation component (see figure 5). There are almost no roll-violations found in pure pitch conditions; the non-varied roll dimension never exceeded 2.5°. This fact stands in contrast to the pure roll conditions, where pitch-violations occurred up to 10° (distortions in the non-varied pitch dimension!). As an example, for 1 subject, figure 4 displays the roll- and pitch-violations in the roll (a and b) and pitch (c and d) tilting conditions; compare fig. 4b with 4c!
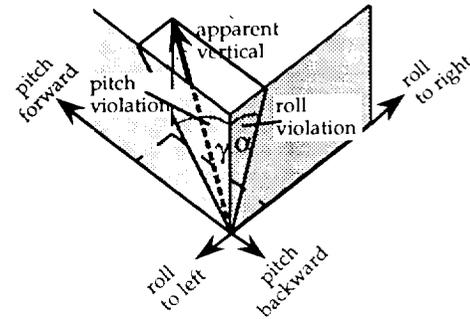
Figure 5
Two components of spatial distorsions:
The pitch violation and the roll violation of
orthogonality (solid line).

DISCUSSION AND CONCLUSIONS

As shown before, the largest deviations from the objective vertical and horizontal are found in body tilts over 90° which are of coarse quite unusual in everyday life. This fact fits nice to results published by Ellis, Kim, Tyler, McGreevy & Stark (1985) and Ellis, Tyler, Kim & Stark (1991), who show in three dimensional tracking experiments that the worst performance is found at 125° misalignment between display and control axes. As discussed there, this might be caused by the mental rotation of space.

The present paper focuses on the *angular distortion* of the perceptual space, regardless of the *extent* or the *quality* to which the perceptual system performs space transformations. Although large interindividual differences were found, the apparent space of all our subjects cannot be considered to be orthogonal. One might conceive that the two slightly different tasks, setting a ring horizontal versus setting a line vertical, could be responsible for the distortions, but, in fact none of the tasks was solved systematically better.

Other explanations could be found like e.g. the anatomy of the vestibular organ or, as proposed by Pellionisz & Llinás (1980) and Pellionisz (1987), the non-orthogonal representation of the 3D space in our brain. "Neurobiological evidence shows, ... that the simplest approach (Cartesian coordinate systems erecting spaces with Euclidean geometries) is untenable for natural systems such as the brain" (Pellionisz, 1991). This would imply that, under conditions of unusual body positions, our perceptual system is not able to reconstruct stored spatial data properly. An other set of experiments with a slightly different setting and body tilts with *combined* pitch and roll angles (Bucher, Mast & Bischof, 1991) confirmed these results.

Certainly, the results are partly due to the artificial experimental environment which does not provide any 3-D objects with familiar angular relations. We probably can be sure that e.g. a presented cube still would be recognized as a cube even if we were tilted 150° sideways. However, the data does imply that a subject experiencing the gravitational force not along the body axis can no longer be expected to estimate angles correctly. Since large interindividual differences were found, it might be necessary to calculate individual distortion matrices to describe angular properties of perceptual space and use them to

579

create virtual environments. An attempt to extract the non-orthogonal portion of the space transformation performed by the CNS is presented in Bucher et al. (1991). An alternative to deal with this problem is to provide an appropriate artificial frame of reference on the visual channel "forcing" the brain to a more orthogonal perception.

Generally the visual display format has a large effect on spatial perception. One has specially to take care of this fact when using graphic displays as planning tools. Ellis, McGreevy and Hitchcock (1987) and Ellis, Kim, Tyler, McGreevy & Stark (1985) have clearly shown the benefits of graphical 3-D space information in an air traffic avoidance experiment. Still it might have to be expected that body tilts affect these very same tasks. Therefore we have to be careful in using absolute angles as analog information in physical environments which are likely to be tilted away from the upright as e.g. high performance jet cockpits are.

The errors in depth perception in pure roll conditions might be due to a vestigial compensatory mechanism, the ocular counterrolling: when turning our head sideways our eyeballs try, by counterrolling around their visual axes, to compensate although never matching more than about 10% of the tilt. This causes a vertical shift of the retinal images relative to each other which could be responsible for the observed failure in depth perception. Experiments to clarify this matter are in progress.

Concerning further investigations in spatial perception, this non-orthogonality means that we are to measure all three perceptual axes rather than only the vertical or the horizontal, whenever we want to learn about it under tilted body conditions or in micro- and hypergravity conditions. And, even for experiments with pure roll body tilts we should provide a device to set the apparent direction which allows as well manipulations in pitch direction.

The original motivation for the study was a system analytical approach to the optic-vestibular interaction. In a *descriptive* approach we have pointed out here some important consequences for further analysis of perceptual space properties and implications for virtual environments.

ACKNOWLEDGMENTS

REFERENCES

Bischof, N. (1974). Optic-Vestibular Orientation to the Vertical. In H.H. Kornhuber (Ed.). Handbook of Sensory Physiology: Vol. VI/2. Vestibular System (pp. 155-190). Heidelberg: Springer.

Bucher, U. (1988). Orthogonalität des subjektiven Wahrnehmungsraumes unter Ausschluss visueller Stimuli zur Raumorientierung. Unpublished doctoral dissertation, University of Zurich.

Bucher, U., Heitger, F., Mast, F., Bischof, N. (1990). A novel automatic procedure for measuring ocular counterrolling. Behavior Research Methods, Instruments, & Computers 22 (5), 433-439.

Bucher, U. J., Mast, F., Bischof, N. (1991). The non-orthogonality of subjective perceptual space. Manuscript in preparation.

Ellis, S. R., McGreevy, M. W. (1987). Perspective traffic display format and airline pilot traffic avoindance. Human Factors, 28, 439-456.

Ellis, S. R., Tyler, M., Kim, W. S., McGreevy, M. W., Stark, L. (1985). Visual enhancements for perspective displays: Perspective parameters. Proceedings of the International Conference on Systems Man and Cybernetics. IEEE Catalog # 85CH2253-3, 815-818.

Ellis, S. R., Tyler, M., Kim, W. S. and Stark, L. (1991). Three dimensional tracking with misalignment between display and control axes. SAE International Conference on Environmental System, July 15-18, San Francisco, CA.

Pellionisz, A. (1987). Tensor network theory of the central nervous system. Encyclopaedia of neuroscience (eg. G. Adelman), Birkhauser, 1196-1198.

Pellionisz, A. (1991). Nature's Geometry and how it may be represented in the brain. Abstract of the paper presented at: NASA Ames Research Center, Human Factors Research Division.

Pellionisz, A., Llinás, R. (1980). Tensorial approach to the geometry of brain function: Cerebellar coordination via metric tensor. Neuroscience , 5, 1125-1136.

C-3

An Intelligent Control and Virtual Display System
for Evolutionary Space Station Workstation Design

by

Xin Feng and Russell J. Niederjohn

Advanced Control Technology Laboratory
Department of Electrical and Computer Engineering
Marquette University
Milwaukee, WI 53233, (414)288-6820

Michael W. McGreevy

Human Interface Research Branch
NASA Ames Research Center
Moffett Field, CA 94035, (415)464-5784

ABSTRACT

Research and development of the Advanced Display and Computer-Augmented Control System (ADCACS) for the space station Body-Ported Cupola Virtual Workstation (BP/VCWS) have been pursued at the Marquette University Advanced Control Technology (ACT) Laboratory. This project explores the potential applications of body-ported virtual display and intelligent control technology for the human-system interfacing applications in space station environment. The new system is designed to enable crew members to control and monitor a variety of space operations with greater flexibility and efficiency than existing fixed consoles.

The technologies being investigated include helmet-mounted virtual displays, voice and special command input devices, and microprocessor-based intelligent controllers. Several research topics, such as human factors, decision support expert systems, and wide field of view, color displays are being addressed. A prototype integrated emulator is in development. Our study has demonstrated the significant advantages of this uniquely integrated display and control system, and its feasibility for human-system interfacing applications in the space station command and control environment.

1. Introduction

The NASA Space Station Freedom Cupola workstations will play a key role in supporting crew to control and monitor many critical operations and systems in their missions. Examples include station manipulators, flight

telerobotic servicer, station mobile transporter, control of external video and lights, etc. To support these functions, the cupola workstation must support direct or video observations of operations of payloads, docking and berthing, remote inspection of any external point on space station structure, EVA operations, and other operations. Clearly, the performance of display and human-system interface technology in the space station cupola environment directly affect the efficiency and productivity of the space mission, and the crew safety. The cupola workstation must provide a human-system interface for task control and monitoring which accommodates both direct viewing through any of the cupola windows, and remote viewing through any of the space station's video cameras, including those on telerobots, on remotely controllable vehicles, and on the station itself.

The multi-directional viewing requirements indicate that there is a potential for conflict between the spatial orientation of a fixed, or "hard console" cupola workstation and the directions of external tasks. A fixed workstation favors a single direction, while disfavoring other directions. In addition, the need to divide attention between external views and panel mounted displays will create problems for the crew operator. Further disadvantages of permanently installed workstations in the cupola include a potential conflict with the requirement that cupola windows must be replaceable on orbit, since the fixed displays are likely to abut, or even block cupola windows. Currently specified workstations will require multiple displays for each workstation, with a penalty of weight, volume, and power.

The fast growing virtual reality and head-mounted display technologies offer an attractive alternative to the fixed console displays. They are portable, easy to install and

easy to maintain. In addition, head-mounted display devices and their control processors are compact, light-weight, and low power. This new technology enables the cupola workstation to offer multiple displays, whose visibility depends on head position and orientation tracking. Moreover, the possible applications of state-of-the-art microcomputer hardware and intelligent control technologies in the cupola workstation design open a promising approach which allows the keyboard, control input devices, and audio input/output devices to be body-ported by the crew operator. With all the new ideas described above, a new design concept, called "Body-Ported, Virtual Cupola Workstation (BP/VCWS)", can be formulated.

Use of head-mounted virtual display and intelligent control technologies in the cupola workstation design, however, create new problems that should be solved before physical construction of the BP/VCWS. First, this newly proposed cupola workstation design involves advanced technologies in multi-disciplines, including a helmet-mounted virtual display device, voice and special command input devices, and microprocessor-based intelligent controllers. Several research topics, such as human factors, decision support expert systems, and wide field of view, color displays must also be addressed. More-over, the effectiveness and user acceptance of such a device is another key issue.

Most head-mounted display technologies, which are currently available for application to the space station has been in use for aviation. Representative systems devel-oped for use in combat rotorcraft or fighter aircraft include the Honeywell Integrated Helmet and display Sight System (HADSS). Two systems developed for use as visual systems for flight simulators are the Visually Coupled Airborne Systems Simulator (VCASS), developed at Wright-Patterson AFB, OH, and the Fiber-Optic Helmet Mounted Display (FOHMD), developed at Williams AFB, AZ. All of these systems have been developed for specific application environments which are very different from that of the space station. For example, there is no need of a crash helmet during missions on board the space station, so the integration of a display into a helmet is of little use. The tradeoffs required for such integration are inappropriate for application to space station. The simulator displays, on the other hand, are very bulky and would be unacceptable as flight equipment.

NASA has two head-mounted display activities that are relevant to the current Advanced Development effort. One project is the development at JSC of prototype displays for use inside EVA helmets. The EVA helmet display has considerable power constraints which limit the design options. Further, the information to be presented on the

EVA display is of far less complexity than the potential IVA display. Overall, the design goals of the IVA system would not be met by the EVA system. The other NASA project is the development at ARC of low-cost, wide field-of-view head-mounted displays for laboratory research in Virtual Environment Display (VIVED). This system does not provide the see-through display capability which will be required on the space station. Further, the ARC system provides a very wide instantaneous visual field-of-view at the expense of resolution, which is the opposite of the design tradeoff needed for the space station display.

The research and development effort for such a BP/VCWS system has already been pursued at the Marquette University Advanced Control Technology (ACT) Laborato-ry. The research project, "Advanced Display and Comput-er-Augmented Control System (ADCACS)", is sponsored by NASA Ames research center and includes the partici-pation of Astronautics Corporation of America. The project motivates the exploration of the potential applica-tions for the virtual display and intelligent control technolo-gies especially in the space station environment. The new system will be designed to enable crew members to control and monitor a variety of space operations with greater flexibility and efficiency than existing fixed con-soles.

The technologies being investigated include helmet-mounted virtual displays, voice and special command input devices, and microprocessor-based intelligent controllers. Several research topics, such as human factors, decision support expert systems, and wide field of view, color displays are being addressed. A prototype integrated emulator is in development. Our study has demonstrated the significant advantages of this uniquely integrated display and control system, and its feasibility for human-system interfacing applications in the space station command and control environment.

2.   ADCACS Project Overview

The Advanced Display and Computer Augmented Control System (ADCACS) was proposed as a three-year systems engineering project that is devoted to the research and development of such an advanced human-system interface device, i.e., the Body-Ported Virtual Cupola Workstation (BP/VCWS). This project takes advantage of other closely related on-going projects that are critical and very important to the success of the project.

Objectives

Specifically, the overall objectives of the ADCACS project are:

583

1) to demonstrate the feasibility of the Body-Ported Cupola Workstations (BP/VCWS) meeting the needs of the baseline and evolutionary Space Station Freedom (SSF) Cupola Workstations;

2) to initiate an engineering approach to construct such a BP/VCWS, and through evolution, to replace existing "hard console" cupola workstation design;

3) to develop a flight-qualifiable prototype BP/VCWS, and to significantly advance the general state-of-the-art in workstation technology including virtual reality" and intelligent Control.

Evolutionary Approach

The above goals were planned to be accomplished in the three-year project that was scheduled in three phases. The Phase one research, started July 1, 1990 and ended June 30, 1991, has been carried out at the Marquette University Advanced Control Technology (ACT) Laboratory, with participation of Astronautics Corporation of America (ACA). During the Phase One research, it was necessary and important to establish an evolutionary approach toward the physical construction of the flight-qualifiable BP/VCWS, starting from the requirements and feasibility analysis to the construction and evolution of on-ground emulators. A variety of factors were taken into account.

The evolutionary steps can be described as:

1) to define requirements and configurations for the BP/VCWS that are consistent with crew user requirements and Space Station Program constraints;

2) to study the relevant technologies, that will be applied to the proposed body-ported virtual display and knowledge-based intelligent control system, through literature search and laboratory experimental investigations; and

3) to build the prototype emulator, and later build the Full-Function Emulator (FFE), based on the prototype emulator for further studies of technology readiness of flight-qualifiable BP/VCWS.

3. Description of the Prototype Emulator System

3.1 Overview of the Virtual Workstation Emulator

A prototype emulator for the virtual display and intelligent system has been designed and built during the first year.

This emulator environment consists of an IBM-386 compatible personal computer (PC1), an IBM-286 compatible personal computer (PC2), and a Silicon Graphics Personal IRIS workstation (IRIS). The IRIS, model 4D/35TG has a 35MHZ CPU. This architecture is shown in Figure 1.
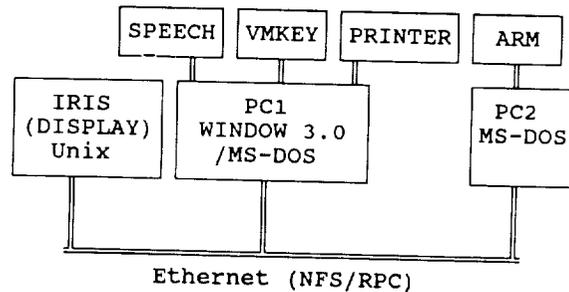

Figure 1   Overview of the Prototype Emulator Environment

As shown in Figure 1, the PC1 is dedicated to the intelligent control and voice command input experiments, and PC2 is used for acquiring and processing data from the fingers/hand/arm position sensing input devices. The IRIS is used for the graphic generation of the virtual scenario elements to be manipulated. The entire system is networked together using an Ethernet LAN. This prototype emulator can perform the following experiments:

1) virtual tele-robot arm/hands display and control;
2) voice command input;
3) force-based arm and position-based master command input and control;
4) knowledge-base intelligent control and decision support system testing and implementation.

3.2 Networking

For inter-process communication over the Ethernet LAN, the "Sockets" protocol was chosen for the ADCACS application. RPC (remote procedure call) and NFS (network file system) were also evaluated as options to sockets.

Internet Domain Stream Sockets was chosen for its simplicity in satisfying the ADCACS project needs. With proper programming, all subsystems can process data concurrently. Any subsystem may be a client or a server. This increases the flexibility of the entire emulator. Minimal overhead is required since sockets is a fairly low level communications technique.

3.3 Graphics Generation

Graphical generations for virtual environment and virtual tele-robot are performed on the Cilison Graphics IRIS workstaiton. This machine is fast enough to generate the real-time (no perceivable delay) images necessary for the present research. It is anticipated that as the complexity of the control scenarios increases, the graphical image generation power will have to be increased.

The primary graphics program in use at the present time is "armbox.c". It uses the Silicon Graphics standard graphics library, GL. GL is a set of C callable subroutines. The subset of routines being used in "armbox.c" includes: 1) a variety of viewing windows; 2) a single infinite light source (multiple infinite and local lighting source models are available for experimentation and require increased processing time); 3) RGB mode and Gouraud shading for realistic shading with fewer polygons as compared with flat shading; 4) double buffering to synchronize rendering with the screen so smooth motion exists with no flickering; 5) backface removal and z-buffering to decrease rendering time for hidden surface removal; 6) viewing matrix manipulation and stack routines to easily and quickly position graphical elements; 7) queues for internal events and standard input devices; and 8) simple 3 dimensional polygon rendering routines.

The program routine "armbox.c" renders a picture of a right arm, including a hand and fingers using positioning data received over the network. It also has techniques for drawing boxes in any location and orientation. For iteration experimentation of the graphics portion of the emulator, a second process may be run on the Personal IRIS which emulates a microcomputer sending data to the IRIS.

For the networking setup, the IRIS is the client and PC2 is the server. The main graphic loop requests positioning data from the server. It then immediately renders a picture using initial data. After rendering the picture, it reads the data sent from the server and loops. The second time through the loop it again requests data from the server and immediately renders a picture, this time with the previous data. Each time though the loop the picture is rendered using the previous data. This technique allows the graphic generation to be performed in parallel with the acquiring and preprocessing of the positioning data.

The screen is refreshed at 60 Hz so there is no perceivable flicker. The image update time appears to be slower than 60 Hz. Until the next image is generated, the IRIS simply re-displays the same image. For the same 3 dimensional model, the image generation time will differ depending on the depicted position of the virtual arm.

Collision detection has been successfully completed for boxes in any orientation. In addition to drawing an arm, several boxes are also drawn within arm's reach. The location of the tip of each finger is compared to the location of a box resulting in a Boolean array of collision locations. Additional points on the fingers, hand, or arm may be identified as important for collision detection and added to the Boolean array. This may be done for any number of boxes.

Once the collision detection data is acquired, decisions will be based on the data. These decisions could include coloring the appropriate box, limiting motion of the moving object (a finger presently), or grabbing and moving one of the boxes. If these decisions are simple and need to be made quickly, they could be performed internal to the IRIS. If they are complex and could be made with a longer time delay, it may be appropriate to off-load the decision process to another processor on the network to preserve processing time on the IRIS for complex graphical generation. The decision processor could be running a knowledge-based software package.

The present collision information is used to draw the appropriate box with a color that is dependent on which finger is touching the box. This gives the operator clear visual feedback.

Also, a virtual space cupola environment model has been developed on the IRIS for environmental and space operation monitoring purposes. This model is connected to the knowledge-based system on PC1 through Ethernet Link.

3.4 Master Command Input Devices

Force-Based Master Arm

A force-based master arm controls the prototype emulator environment. The arm controller utilizes 5 degrees-of-freedom on the arm (i.e. shoulder flexion/extension, shoulder abduction/adduction, upper arm rotation, elbow flexion/extension, and forearm rotation) with near future goals of wrist flexion/extension and wrist abduction/adduction. The current configuration consists of the force-based arm as an input device which sends an analog signal to an A/D board and the interpretation software in a 286-based microcomputer (PC@). This computer reads a digital signal from the 5 degrees-of-freedom on the arm and maps it to a range of motion for the virtual arm. This information is then sent across a

distributed network to the graphics workstation via Internet Domain Stream Sockets. This data is read by the graphics terminal and is used to adjust the position of the virtual human arm that is graphically portrayed.

The interpretation that occurs at PC2 consists of scaling the digital output of each channel to a range of motion. This is accomplished through the utilization of a physiological data base called Ergobase by Biomechanics Corporation of America. This data base contains the range of motion for each degree-of-freedom that is addressed by the force-based arm. Ergobase contains the ranges of motion in degrees for the above specified movements for the central 95 percentile of the population. By linearly transforming the digital output of the force-based sensor to an angle degree that is within the human range of motion for the specific joint, a truly anthropomorphic graphical arm in the virtual world results.

The prototype emulator has been created with optimum anthropomorphic and human factor considerations at every step of its design. The design of virtual scenario presentation will be dependent on calculations and analysis of optimal human range of motion.

Position-Based Master (PBM) Hand

The position based master (PBM) hand has been successfully added to the fingers/hand/arm position data processing microcomputer to complement the force-based arm. Each joint of the PBM hand has an associated potentiometer (pot). The resulting voltage signal from each pot (some limited range within 0 to 5 volts) is sent through a two stage operational amplifier circuit. The first stage isolates the signal from the rest of the circuit. This is useful since it essentially eliminates any effects of unmatched pot and wire resistances in the PBM. The second stage has two adjustable pots. One is for gain and the other is for level shifting so that the resulting voltage ranges from 0 to 5 volts. The output of the second stage is properly conditioned for the full range of the analog-to-digital converter.

3.5   Knowledge-Based Intelligent Control

One 386-based personal computer is dedicated for experimental investigations of knowledge-based intelligent display system control and the decision support systems. A microcomputer-based speech recognition system, Voice Master Key produced by Convox, Inc., was successfully installed at the beginning of the ADCACS project. A 386 microcomputer is designated for intelligent expert system control to the speech recognition system. Based on PC1 and communication facilities through LAN, the expert

system shell, called "Nexpert Object", is used for the following two intelligent control experiments:

1)   A Prototype KBS Control System (see Figure 2 for details), and
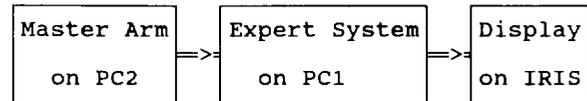

Figure 2a: Control of Data Flow from Master Arm to IRIS
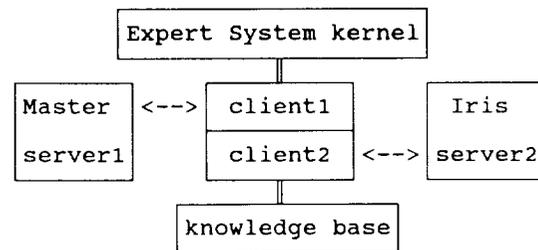

Figure 2b: Control Signal Diagram
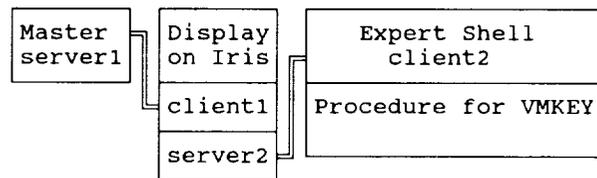
Figure 2c: Display Synchronization Diagram

2)   A simple control system that is based on a time-sequential control strategy, which uses several procedures implemented by one process in PC1. The calling out function of Expert will be used, which refers to the ability of an expert system application to call user written functions or library routings. The knowledge-based Expert System Shell was used to infer the production rules in the knowledge base for data validity checking and intelligent controlling. Also the data transform and inter-procedure communications were implemented under the Knowledge-Based System.

3.6 Voice Command Input Devices

The Voice Master Key (VMKey)

VMKey (Version 2.0) is a voice control device made by CONVOX INC., which can be used to train and control computers to accept commands in simple spoken English or any other languages. VMKey comes in two parts: the hardware "ear" and the software "brain". The "ear" is a combination headset earphone/microphone plus a plus-in card. This device allows the user to run an application program with voice commands.

Voice Command Process Control

The following voice command input functions have been implemented:

1) VMKEY was used to accept the control interface signals;

2) VMKey has been Integrated into the Expert System; and

3) The SPEECH Input Device has been employed to demonstrate the voice warning functions.

4. Conclusions

The design and development of the prototype emulator implements the first step of the evolutionary approach towards the construction of the proposed BP/VCWS system. Since this emulator is only a prototype, it has limited functions. Future work with the goal of improving the emulator includes integration of wide-field-of-view optical devices, implementation of stereoscopic display devices, and simulation of operational warning and caution systems. However, this prototype emulator establishes a solid foundation for the construction of the full function emulator for more intensive technical investigation of the proposed BP/VCWS.

References

1. Banks, S. and Lizza, C., "Pilot's Associate: A Cooperative, Knowledge-Based System Application," IEEE Expert, Vol. 6, No. 3, June, 1991, pp. 18-29.

2. Fisher, S., McGreevy, M., Humphries, J., Robinett, W., "Virtual Environment Display System," ACM Workshop on Interactive 3-D Graphics, 1986, UNC, Chapel Hill, NC.

3. Fisher, S., "Telepresense Master-Glove Controller for Dexterous Robotic End-Effectors," SPIE Vol. 726, Intelligent Robots and Computer Vision, 1986, pp. 396-401.

4. Fisher, S., McGreevy, M., Humphries, J., Robinett, W., "Virtual Workstation: A Multimodel, Stereoscopic Display Environment," SPIE Vol. 726, Intelligent Robots and Computer Vision, 1986, pp. 517-522.

5. Furness, T., Designing in Virtual Space, chapter in System Design, W.B. Rouse & K.R. Boff, eds., North Holand, 1987.

6. Furness, T., "Helmet-Display Reading Performance During Whole-Body Vibration," Presentation, 52nd Annual Meeting of Aerospace Medical Association, San Antonio, TX, May, 4-7, 1981.

7. Furness, T., "The Super-Cockpit and its Human Factors Challenges," Proceedings of the Human Factors Society 30th Annual Meeting, pp. 48-52, 1986.

8. Furness, T., "Super Cockpit Amplifies Pilot's Senses and Actions," Government Computer News, Aug. 15, 1988.

9. Mark, W., and Simpson, R., "Knowledge-Based Systems: An Overview," IEEE Expert, Vol. 6, No. 3, June, 1991, pp. 12-17.

10. McGreevy, M., "Keynote Speech: Personal Simulators and Planetary Exploration," CHI'89, May, 1989.

11. McGreevy, M., " Virtual Workstation Overview, FY88," Technical Report, NASA Ames Research Center, CA, 1989.

12. "Space Station Integrated Workstation Requirements and Definition Document," NASA Space Station Document, Lyndon Johnson Space Center, January, 1989.

13. Summut, C. and Michie, D., "Controlling a Black Box Simulation of a Space Craft," AI Magazine, Vol. 12, No. 1, Spring 1991, pp. 56-63.

14. Zimmerman, T., Lanier, J., Blanchard, C., Bryson, S., Harvill, Y., "A Hand Gesture Interface Device,: ACM SIGCHI/GI, 1987, pp. 189-192.

|

Session H6: SPACE PERCEPTION AND PHYSIOLOGY - II

Session Chair: Dr. Barbara Stegmann

TRACKING PERFORMANCE WITH TWO BREATHING OXYGEN CONCENTRATIONS AFTER HIGH ALTITUDE RAPID DECOMPRESSION

Thomas E. Nesthus, Ph.D.
KRUG Life Sciences
San Antonio Division
P.O. Box 790644
San Antonio, TX. 78279

Samuel G. Schiflett, Ph.D.
and Carolyn J. Oakley
AL/CFTO
Brooks AFB, TX.
78235-5000

ABSTRACT

Current military aircraft Liquid Oxygen (LOX) systems supply 99.5% gaseous Aviator's Breathing Oxygen (ABO) to aircrew. Newer Molecular Sieve Oxygen Generation Systems (MSOGS) supply breathing gas concentrations of 93-95% oxygen. This study compared the margin of hypoxia protection afforded by ABO and MSOGS breathing gas after a 5 psi differential rapid decompression (RD) in a hypobaric research chamber. The barometric pressures equivalent to the altitudes of 46,000, 52,000, 56,000, and 60,000 ft were achieved from respective base altitudes in 1-1.5 s decompressions. During each exposure subjects remained at the simulated peak altitude breathing either 100% or 94% O_2 with positive pressure for 60 s, followed by a rapid descent to 40,000 ft. Subjects used the Tactical Life Support System (TLSS) for high altitude protection. Subcritical tracking task performance on the Performance Evaluation Device (PED) provided psychomotor test measures. Overall tracking task performance results showed no differences between the MSOGS breathing oxygen concentration of 94% and ABO. Significant RMS error differences were found between the ground level and base altitude trials compared to peak altitude trials. The high positive breathing pressures occurring at the peak altitudes explained the differences. Considered with the physiologic data, an acceptable degree of hypoxia protection was met with both oxygen concentrations using TLSS at altitudes <60,000 ft for <60 s durations.

INTRODUCTION

In both the US Navy and the US Air Force, there is increasing interest in Molecular Sieve Oxygen Generation Systems (MSOGS) for their logistic and reliability advantages when compared to liquid oxygen supplied aircraft breathing systems. A limitation in the maximum oxygen concentration attainable with MSOGS, however, has motivated USN and USAF development communities to establish laboratory evidence of the acceptability of using reduced breathing oxygen throughout the altitude envelope of current aircraft oxygen systems.

Based upon a fairly well developed theory of respiratory gas exchange at altitude, our team of researchers concluded that there was no reason to expect adverse effects of MSOGS oxygen concentrations at normal cabin pressures. However, after a rapid loss of cabin pressure while flying at emergency ceiling altitudes needed further investigation. Especially, if a reduction of oxygen concentration is expected in the breathing gas supplied to the aircrew. We therefore, incorporated a rapid decompression (RD) profile in our study.

The first phase of research employed the current production oxygen system including: the CRU-73 dilution-demand breathing regulator and it's oxygen delivery/breathing pressure schedule; the MBU 12/P oxygen mask and HGU 55-P helmet. The RD profile was across a 5 psi differential, from 20,000 to 50,000 ft, and remained at peak altitude for 60 s. Results of this phase of research were reported elsewhere (Bomar, et. al, 1988; Holden, et. al, 1987; Nesthus, et. al, 1988; Nesthus and Schiflett, 1989; Wright, et. al, 1988; Wright, et. al, 1990).

During the second phase of study we used a developmental life support system designed to improve high altitude and high acceleration protection. The Tactical Life Support System (TLSS) included a modified CRU-73/TLSS dilution-demand oxygen regulator with an adjusted oxygen delivery and breathing pressure schedule. Also, a TLSS helmet, mask, and counterpressure jerkin-vest

PRECEDING PAGE BLANK NOT FILMED

system was used to allow breathing gas delivery at much higher positive pressures needed for high altitude protection.

Our altitude profile simulated loss of cabin pressure while flying at various potential emergency flight ceilings. The profile incorporated a 5 psi differential RD similar to Phase I research but we included 4 different base-to-peak simulated altitudes seen in Table I.

Both phases of study were conducted in the hypobaric research chambers at the USAF School of Aerospace Medicine (USAFSAM), Brooks AFB, Texas.

Table I: Four base-to-peak, 5 psi differential rapid decompression profile pressures and simulated altitudes.

Pressure (torr)		Altitude (ft)	
Base	Peak	Base	Peak
364.4 - 105.9	=	19,000 - 46,000	
340.0 - 79.5	=	20,800 - 52,000	
321.3 - 65.7	=	22,000 - 56,000	
307.9 - 54.2	=	23,000 - 60,000	

METHOD

Our subject population was comprised of 17 chamber-qualified active duty male volunteers from the USAFSAM Altitude Panel. The voluntary fully informed consent of the subjects used in this research was obtained as required by AFR 169-3.

In addition to measuring a number of physiologic parameters, discussed in detail in the Phase I research references, a computer-based unstable tracking task from the Performance Evaluation Device (PED) provided two psychomotor measures (Systems Research Laboratory, 1987). The tracking tasks' instability was based on an algorithm similar to that of the subcritical tracking task (Jex, 1967). RMS offset-from-center error and the number of boundary hits were the primary measures of tracking performance. Subjects were trained to perform the task while inside the chamber environment wearing the TLSS ensemble with most of the physical distractors in place. Sessions with high positive breathing pressures were also included.

Figure 1 shows a generic altitude profile and time line for one experimental RD session.

During a 1 hr 100% O_2 prebreathe for decompression sickness prevention, one performance task warm-up and trials 1 and 2 were conducted. An ear and sinus check and an abdominal gas check were made before holding at the base altitude. Pre-RD physiological

recordings and trials 3 and 4 were completed. Prior to the RD, the breathing gas mixture was switched from 100% oxygen to a pre-RD mixture of O_2 representative of the CRU-73's scheduled dilution mixture for each particular base altitude. Subjects breathed this mixture for 2-3 minutes for pulmonary equilibration. The base altitude breathing gas mixture and the peak altitude oxygen condition for each experimental trial was unknown to the subject. After a final "ready" was communicated, the subject was cautioned to breath normally. Then the hypobaric chamber was rapidly decompressed (approximately 1 s) to a simulated peak altitude of either 46,000 ft, 52,000 ft, 56,000 ft, or 60,000 ft. The positive breathing pressure at 46,000 ft, irrespective of the O_2 condition, was 50 mmHg at the mask. Positive breathing pressure at the remaining peak altitudes was 70 mmHg. The subject, initiating the "Peak" performance task trial ten seconds after the RD, remained at that altitude for 50 s more, whereupon the chamber pressure was increased to a 40,000 ft equivalent (141.18 torr). When the subject completed the unstable tracking task a descent to ground level was made. This procedure was repeated for each O_2 condition and peak altitude.
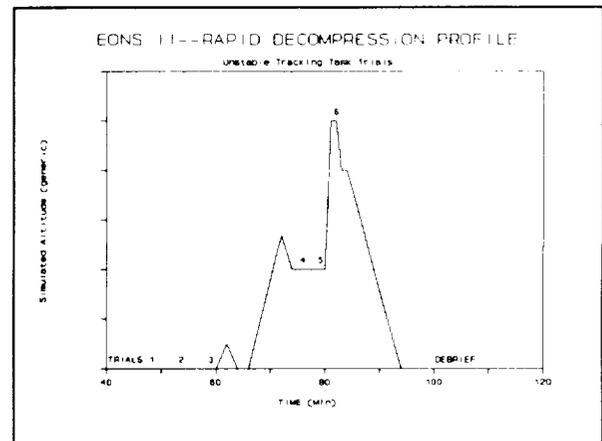

Figure 1: A generic altitude profile for the EONS II rapid decompression study.

A mixed, random and fixed effects design was followed. The fixed effects included: two peak altitude oxygen conditions--100% O_2 and 94% O_2; four peak altitude conditions--46,000, 52,000, 56,000, and 60,000 ft; and three trial levels--Ground, Base, and Peak. Measures analyzed for this report included: Root-Mean-Squared offset from center (RMS) and boundary hits or control losses for unstable tracking performance; and one physiologic parameter, oxyhemoglobin saturation (SaO_2).

591

RESULTS

Our overall design analysis revealed 3-way interactions (O_2-by-Level-by-Peak Altitude) for RMS tracking error, Boundary Hits (BHITS), and SaO_2. These results were anticipated. Separate analyses for O_2 and Peak Altitude were conducted and resulted in predominant Level effects for RMS error and SaO_2. The former was due primarily to the combined effects of positive breathing pressures delivered at the peak altitudes and potential hypoxia. No positive breathing pressure was delivered at ground and base levels. These results can be seen in Figure 2 for the 100% O_2 condition and in Figure 3 for the 94% O_2 condition.
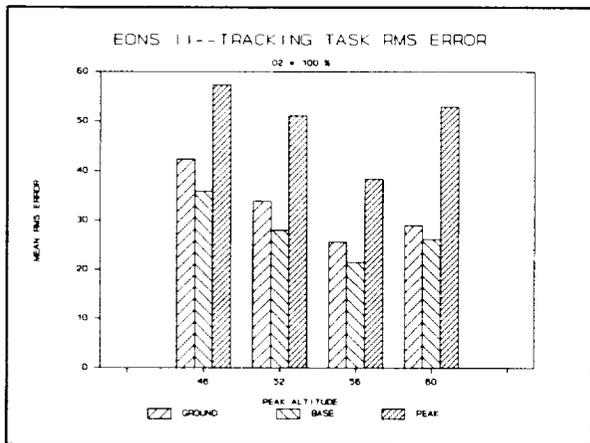
Figure 2: Mean RMS error by Level and Peak Altitude for the 100% O_2 Condition.
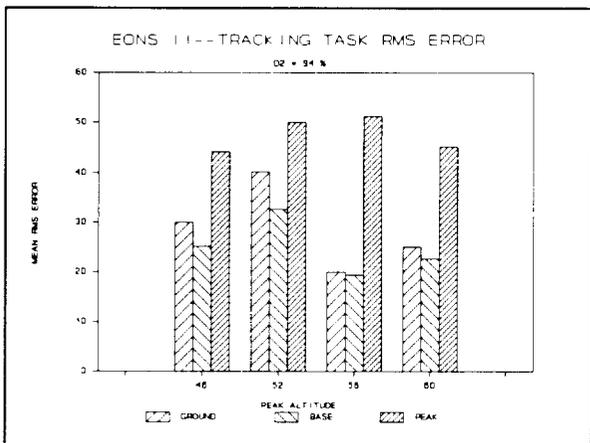
**Figure 3:** Mean RMS error by Level and Peak Altitude for the 94% O_2 Condition.

The Level effect for SaO_2 was primarily due to high oxyhemoglobin saturations which occured while breathing 100% O_2 during the ground and base level trials (prior to the RD) compared to high altitude desaturations which occured at

peak altitudes. This effect is seen in Figure 4.

The Level analysis revealed an O_2-by-Peak Altitude interaction which is clearly seen in Figure 5. Least Square mean t-tests showed that boundary hits for the 94% O_2 condition were greater at 52,000 ft compared to 56,000 or 60,000 ft.
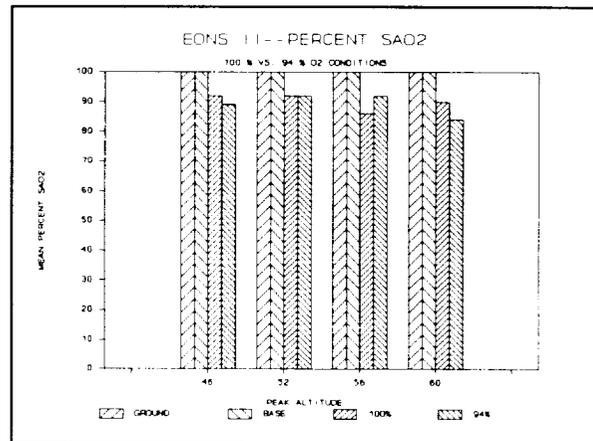

Figure 4: Mean minimum SaO_2 percentage by Level and Peak Altitudes for the 100% and 94% O_2 Conditions.
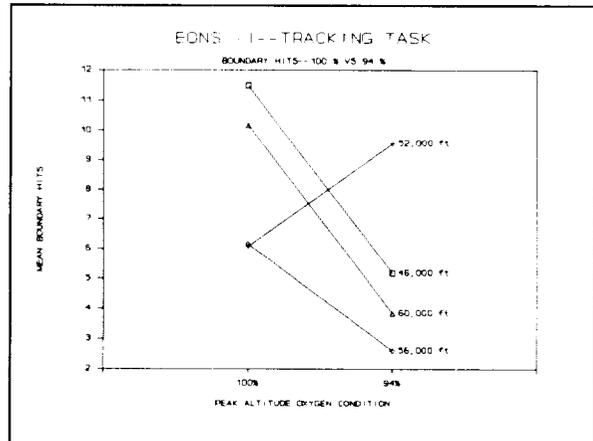
Figure 5: Mean Boundary Hits for O_2 by Peak Altitude interaction

Figures 6 and 7 are examples of additional physiologic data showing 5 s mean $P_{ET}O_2$ values (with +/- standard error) 10 s before and 80 s after RDs to 60,000 ft for the 100% and 94% O_2 conditions, respectively. The figures show a rapid fall in PO_2 at the RD (verticle line in figures) followed by relatively stable values before the descent to 40,000 ft (at time 60 s in figures) as an increase in barometric pressure occurred. The values indicated subjects were exposed to compensatory levels of hypoxia as described in the

592

USAF Physiological Training Pamphlet
(Tables 4-3 and 4-5). Any performance
deficit assumed at this level of hypoxia
was confounded with the positive
breathing pressures at peak altitudes
and were probably diminished by the
transient exposure (i.e., <60 s). The
relatively high SaO₂ values seen in
Figure 4 at peak altitudes may also
reflect the transient nature of the
exposure.
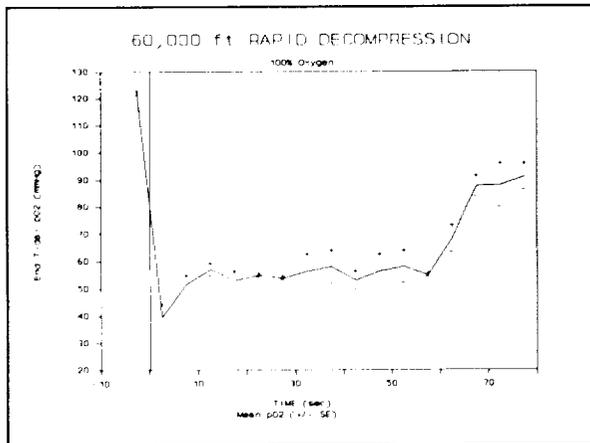
Figure 6: Mean (5 s epoch) End-Tidal pO_2
before and after rapid decompressions to
60,000 ft for the 100% O_2 Condition.
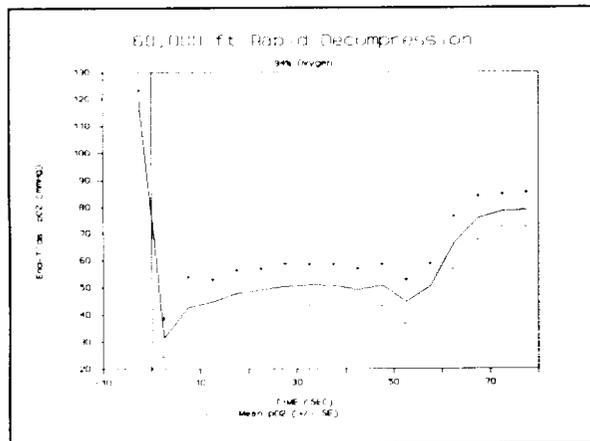
Figure 7: Mean (5 s epoch) End-Tidal pO_2
before and after rapid decompressions to
60,000 ft for the 94% O_2 Condition.

DISCUSSION

We believe decrements found in
tracking performance for the Level
effect were due primarily to the
anticipated effects of high positive
breathing pressures delivered to the
mask for high altitude protection.
Confounded in these data, however, are
potential effects due to the transient,
compensatory level of hypoxia

experienced by the subjects in this
phase of research. We feel the increase
in RMS error was not of a magnitude
which would translate into operational
instability.

The O_2-by-Peak Altitude interaction
for the boundary hits measure, as
displayed in Figure 5, demonstrated the
only evidence of a performance decrement
with the 94% O_2 condition compared to
the 100% O_2 condition. The elevated
mean boundary hits found at 52,000 ft
for the 94% condition were not fully
understood. A thorough investigation of
the data and various post-hoc tests did
not help us explain this effect. No
other performance differences were found
between the 100% and 94% conditions.

CONCLUSIONS

We conclude that unstable tracking
performance was not appreciably
different for the two oxygen conditions
compared. The combined effects of
positive breathing pressure and possible
hypoxia during the peak altitude trials
affected unstable tracking performance
by increasing RMS error. High breathing
pressures were necessary for high
altitude protection and were not present
during the ground or base level trials.
Overall, we believe the TLSS provided an
adequate degree of protection against
hypoxia for both oxygen conditions for
durations less than 60 s at altitudes up
to 60,000 ft as were studied in this
phase of research.

ACKNOWLEDGEMENTS

We wish to express our sincere
gratitude toward key members of the Crew
Systems Branch of the Crew Technology
Division including: Cols John Bomar, Jr.
and Roger Stork, Lt Col Roberta Russell,
Mr Ron Holden, Lts Catherine Wright and
Rob O'Connor for all direction and
support provided to the "cognitive
performance" group during the two phases
of "EONS" research. Additional thanks
are extended to Ed Lee and Irma Baker of
KRUG Life Sciences for their assistance
in running most of the subjects during
this second phase of research.

REFERENCES

Bomar, J.B., Jr., Holden, R.D.,
O'Connor, R.B., Wright, C.S., and
Nesthus, T.E. "Effect of OBOGS Breathing
Gas Concentrations On Crew Performance
At High Altitude" Aviation, Space and
Environmental Medicine, 59, 5, 1988,
Abstract No. 53, p. 472.

Holden, R.D., Bomar, J.B., O'Connor,
R.B., Wright, C.S., and Nesthus, T.E.
"Acceptability of standard USAF

breathing gear at high altitude" In
Proceedings of the 25th Annual SAFE
Symposium Newhall, CA: SAFE Association
1987, pp. 166-170.

Jex, H.R., McDonnell, and Phatak, A.V.
"A 'critical' tracking task for manual
control research" IEEE Transactions on
Human Factors in Electronics. HFE-7,
1966, pp. 138-144.

Nesthus, T.E., Schiflett, S.G., and
Bomar, J.B., Jr. "The Effects of
Different Breathing Gas Concentrations
On Task Performance At High Altitude"
Aviation, Space and Environmental
Medicine, 59, 5, 1988, Abstract No. 52,
p. 471.

Nesthus, T.E. and Schiflett, S.G., "The
Effects Of Various Breathing Gas
Mixtures On Cognitive And Psychomotor
Task Performance" Poster at The 33rd
Annual Meeting of the Human Factors
Society, Denver, Colorado 1989.

Systems Research Laboratories
"Performance Evaluation Device (PED)
Operators Manual", Version 2.1. SRL,
Chesapeake Bay Office, California,
Maryland, 1987.

U.S. Department of the Air Force
"Physiological Training Pamphlet" (AF
Pamphlet 160-5), Headquarters U.S. Air
Force, Washington, D.C., 1976.

Wright, C.S., O'Connor, R.B., Holden,
R.D., Nesthus, T.E., and Bomar, J.B.,
Jr. "Heart Rate Response After Rapid
Decompression To 50,000 Feet: The Effect
of 93% Oxygen Breathing Gas" Aviation,
Space and Environmental Medicine, 59, 5,
1988, Abstract No. 54, p.472.

Wright, C.S., Harding, R.M., Bomar,
J.B., Jr. Holden, R.D., and Bauer, D.H.
"Rapid Decompression To 50,000 Feet:
Effect on Heart Rate Response" Aviation,
Space and Environmental Medicine, 61,
1990, pp. 604-608.

SPACE SICKNESS PREDICTORS SUGGEST FLUID SHIFT INVOLVEMENT

AND POSSIBLE COUNTERMEASURES

K.E. Simanonok, E.C. Moseley, and J.B. Charles

NASA/Johnson Space Center, Houston, TX 77058

ABSTRACT

Preflight data from 64 first-time Shuttle crew-members were examined retrospectively to predict space sickness severity (NONE, MILD, MODERATE, or SEVERE) by discriminant analysis. From nine input variables relating to fluid, electrolyte, and cardiovascular status, eight variables were chosen by discriminant analysis that correctly predicted space sickness severity with 59% success by one method of cross-validation on the original sample and 67% by another method. The eight variables in order of their importance for predicting space sickness severity are sitting systolic blood pressure, serum uric acid, calculated blood volume, serum phosphate, urine osmolality, environmental temperature at the launch site, red cell count, and serum chloride. These results suggest the presence of predisposing physiologic factors to space sickness that implicate a fluid shift etiology. Addition of a tenth input variable, hours spent in the Weightless Environment Training Facility (WETF), improved the prediction of space sickness severity to 66% success by the first method of cross-validation on the original sample and to 71% by the second method. The data suggest that WETF training may reduce space sickness severity. Astronauts while in WETF training are not oriented by gravity but remain physiologically susceptible to it, causing cephalad shifts of fluid when they are head down. As a result, they may physiologically adapt to the fluid shifts by reducing their blood volume, which also occurs in weightlessness. If fluid shifts contribute to space sickness, preadaptation of the circulation by preflight blood volume reduction may counteract fluid shifts in weightlessness and help to alleviate space sickness.

INTRODUCTION

Space sickness is an operationally significant problem that occurs to varying degrees of severity in about two-thirds of Shuttle astronauts (Davis et al., 1988). Besides being a nuisance, mission performance can be negatively impacted when astronauts are severely space sick, and safety is an important issue. An emergency Shuttle landing early in a mission could be endangered if the pilot and commander were spacesick (Vanderploeg et al., 1985). The risk of fatality from vomiting in a spacesuit is presently minimized by an operational policy prohibiting extravehicular activity (EVA) during the first three mission days when space sickness is most likely to occur. Space sickness typically begins about an hour after orbital insertion, reaches a peak within 24 to 48 hours, and usually resolves between 30 to 48 hours, although it can persist for 72 hours. The time course is variable, however, with occasional delayed-onset space sickness occurring after two days in space (Thornton et al., 1987). NASA has expended great effort in attempting to understand, predict, and treat space sickness to relieve crew discomfort, increase mission productivity, and enhance safety.

Space sickness did not occur in the early days of spaceflight; apparently the larger cabins in Apollo and later spacecraft permitted enough mobility to exceed vestibular susceptibility thresholds for the induction of space sickness. However, susceptibility to vestibular stimulation is not likely to be the sole etiologic factor in space sickness, because preflight tests of astronauts' motion sickness susceptibilities do not correlate significantly with space sickness and are not useful for prediction (Homick et al, 1987). Although the dominant paradigm at this time regards the space sickness problem as a form of motion sickness resulting from novel neurovestibular stimulation in the unique environment of weightlessness (Crampton, 1990), it may be more fruitful to consider space sickness as a binary process. While the dominant paradigm is probably correct in that an essential component of space sickness involves some unaccustomed neurovestibular stimulation, a second important factor may be a decreased threshold of susceptibility to nauseogenic stimuli in general, brought on by physiologic responses to weightlessness. Among the insults upon homeostasis in the early hours and days of a mission which might lower an astronaut's tolerance to provocative vestibular stimuli are some of the effects of fluid shifts.

There is believed to be a dramatic headward fluid redistribution with substantial physiologic responses immediately upon exposure to weightlessness (Greenleaf, 1984). This simple picture is complicated somewhat by the fact that fluid shifts probably begin while astronauts wait in the semi-supine position before launch (Lathers, 1989). Regardless, space sickness follows closely upon the time course of both these fluid shifts and the major physiological perturbations they produce. Fluid shifts caused by head-down tilt in bedrest studies sometimes are associated with

595

dizziness and nausea upon head movement, spontaneous nystagmus, and vestibular illusions of tilting and falling (Kakurin et al, 1975). Fluid shifts caused by water immersion have been reported to increase vestibular susceptibility to caloric stimulation so much that sometimes the caloric stimulation had to be stopped (Mitarai et al, 1981).

Several potential mechanisms exist by which fluid shifts may contribute to space sickness, whether or not the fluid shifts begin on the launch pad (Simanonok et al., in review). We have previously shown that nine preflight variables relating to fluid, electrolyte, and cardiovascular status can be used to predict space sickness incidence with about 80% success and space sickness severity with 55% success using two methods of "pseudo" cross-validation on the original sample of 64 subjects (Simanonok et al., in review). In the present study, the same nine predictor variables have been used, with the addition of the astronauts' training time in the Weightless Environment Training Facility (WETF) as a potential predictor of space sickness. This variable was added because Youmans et al. (1987) have reported an inverse relationship between WETF training and space sickness severity.

The WETF is a tank of water measuring 33 by 78 feet and 25 feet deep. It is used to simulate weightlessness for astronauts preparing for EVA; not all astronauts get WETF training. Astronauts who train in the WETF wear full space suits adjusted for neutral buoyancy in the water. Suit pressure is regulated to approximately 4.2 PSI above the water pressure at any given depth, so there is no externally applied hydrostatic gradient on their bodies as occurs in most water immersion studies. However, they are still in a gravity field and they may sometimes assume head-down positions, which would cause headward fluid shifts to occur. It is possible that repeated exposure to fluid shifts in WETF training confers some preflight physiologic adaptation to the astronauts which lessens their physiologic responses to fluid shifts in weightlessness, therefore helping to protect against space sickness if fluid shifts are involved in space sickness etiology. An alternative or adjunct hypothesis is that unusual positional orientations that the astronauts assume relative to the spacecraft mockups in WETF training help to visually adapt them to similar orientations they will experience in weightlessness. Visual adaptation might reduce the impact of the "sensory conflict" conceptions of the dominant paradigm (Crampton, 1990) that may also be an important component of space sickness.

METHODS

Preflight clinical data for 64 first-time Shuttle crewmembers were used in this study. Not all measures were available at the same times preflight, and some measures were made several times preflight. The first available data for each crewmember were used in this order: launch minus 30 days, launch minus 10 days, and launch minus 3 or 2 days. Variables not measured for a given astronaut on any of those occasions were obtained from their annual physicals. Each astronaut's space sickness severity was determined by a NASA flight surgeon according to the ordinal scale shown in Table 1. Space sickness incidence was defined as any occurrence of space sickness

from MILD through SEVERE and called SICK, as opposed to the NONE category. However, the prediction of space sickness incidence was not improved by addition of WETF training over the 80% success obtained by the nine fluid shift variables alone (Simanonok et al., in review), so those results are not reported here. Space sickness severity was defined as the degree of space sickness from NONE through SEVERE. Predicted preflight blood volumes for each astronaut were calculated on the basis of sex, height and weight by the method of Feldschuh and Enson (1977). Free air maximum and minimum temperatures for three days before each Shuttle launch plus launch day at the Shuttle Landing Facility at Kennedy Space Center were obtained from the Landing Support Office at the Johnson Space Center. Means of the minimum four-day temperatures were used in the following analyses as the prediction variable representing environmental temperature at the launch site. The WETF training data we used are a subset of those previously analyzed and reported by Youmans et al. (1987) because we did not have available the other prediction variables for some of the astronauts in their larger WETF sample.

TABLE 1

SPACE SICKNESS SCORING CRITERIA
From Davis et al. (1988)

NONE (0)
NO SIGNS OR SYMPTOMS REPORTED WITH THE EXCEPTION OF MILD TRANSIENT HEADACHE OR MILD DECREASED APPETITE.

MILD (1)
ONE TO SEVERAL SYMPTOMS OF A MILD NATURE; MAY BE TRANSIENT AND ONLY BROUGHT ON AS THE RESULT OF HEAD MOVEMENTS; NO OPERATIONAL IMPACT; MAY INCLUDE SINGLE EPISODE OF RETCHING OR VOMITING; ALL SYMPTOMS RESOLVED IN 36-48 HOURS.

MODERATE (2)
SEVERAL SYMPTOMS OF A RELATIVELY PERSISTENT NATURE THAT MAY WAX AND WANE; LOSS OF APPETITE; GENERAL MALAISE, LETHARGY AND EPIGASTRIC DISCOMFORT MAY BE MOST DOMINANT SYMPTOMS; INCLUDES NO MORE THAN TWO EPISODES OF VOMITING; MINIMAL OPERATIONAL IMPACT, ALL SYMPTOMS RESOLVED IN 72 HOURS.

SEVERE (3)
SEVERAL SYMPTOMS OF A RELATIVELY PERSISTENT NATURE THAT MAY WAX AND WANE; IN ADDITION TO LOSS OF APPETITE AND STOMACH DISCOMFORT, MALAISE AND/OR LETHARGY ARE PRONOUNCED; STRONG DESIRE NOT TO MOVE HEAD; INCLUDES MORE THAN TWO EPISODES OF VOMITING; SIGNIFICANT PERFORMANCE DECREMENT MAY BE APPARENT; SYMPTOMS MAY PERSIST BEYOND 72 HOURS.

Statistical analyses were performed with BMDP (BMDP Statistical Software, Inc., Los Angeles, CA) on either a Digital Equipment Corporation (Westminster, MA) VAX 780 or VAX 3400. Prediction of space sickness severity was performed in this study by discriminant analysis. Chi-square analyses were done to provide a comparison of the prediction success expected due to chance alone

and to estimate the significance level of WETF training on reducing space sickness severity as compared to that which would be expected to occur by chance. Spearman correlation coefficients were done to describe the strengths of relationships among the prediction variables and between the prediction variables and the criterion. Mann-Whitney rank-sum tests were used to test for differences between variables in each of the space sickness groups with their counterparts in the NONE group, and between the two groupings of astronauts that either did or did not receive WETF training. Statistical significance was accepted at p<0.05.

Data Analysis
The discriminant analysis program calculated an analysis of variance in a series of steps, selecting one variable at each step with the greatest F value. A classification function was then derived which reduced the prediction variables into a single weighted composite with appropriate weighting coefficients for separation of the cases into the groups that they belonged. A similar application of discriminant analyses with a more extensive description of rationale and procedures can be found in Reschke et al. (1984). Our previous analyses were performed with nine preflight variables describing some aspects of fluid, electrolyte, and cardiovascular status (Simanonok et al., in review). In that work we initially compared the efficacy of the same nine variables in predicting several different groupings of space sickness incidence and severity, so the same nine variables were used in all cases; i.e., they were "forced" into the discriminant analyses rather than being chosen by the program. In this work we removed the force instruction and allowed the program to choose the variables which best predicted the criterion variable, space sickness score. Variable selection was limited by setting the minimum F-to-enter at 1.5, which provided an approximate F-value in terms of contributing to the predictions at the end of variable selection that was statistically significant at p<0.05.

Two estimates were calculated of the ability of the nine fluid shift variables to predict space sickness severity on an ideal hypothetical cross-validation sample composed of new cases. One type of cross-validation on the original sample termed a "jackknife" cross-validation was calculated, in which each case was removed from the analysis one at a time and new weighting coefficients on the classification functions (keeping the same variables) were computed with the remaining 63 cases. The new classification function was then used to classify each case as it was removed. Second, a subsample cross-validation was performed. Arbitrarily, 40 cases were selected by stratified random sampling and used to create classification functions to predict the remaining 24 cases. In this second method of cross-validation the variables originally chosen as predictors were forced into discriminant analyses to allow an evaluation of their success in predicting space sickness by this cross-validation method.

RESULTS

The nine fluid shift variables used as input in discriminant analyses are shown in relation to space sickness scores in Fig. 1. A similar plot

of the relationship between WETF training time and space sickness severity is shown in Figure 2. Figure 3 depicts the WETF data in a different form, showing the number of astronauts in each space sickness category that did and did not have WETF training, regardless of the number of hours spent. Intercorrelations of the nine original predictor variables, WETF training time, and space sickness severity are shown in Table 2.

Prediction of space sickness severity using nine fluid shift variables as input is shown in Table 4. The program chose eight variables as predictors and used them to create classification functions which could classify the 64 astronauts into the correct space sickness category with 69% success. The eight variables were chosen in this order: sitting systolic blood pressure, serum uric acid, calculated blood volume, serum phosphate, urine osmolality, environmental temperature at the launch site, red cell count, and serum chloride. The jackknife cross-validation success of 59% prediction shows some shrinkage from the classification matrix, a typical phenomenon for cross-validation studies. The results of jackknife cross-validation may best reflect the true probability of correctly predicting new cases of space sickness on the basis of the 64 subjects used in this study. The subsample cross-validation success rate of 67% is more likely to be artifactually high due to chance, because only data from 40 subjects were used to predict space sickness severity in the remaining 24 subjects.

Addition of WETF training time as a tenth variable in another set of discriminant analyses improved the prediction of space sickness severity, shown in Table 5. The program again chose eight of the ten variables as significant predictors of space sickness, this time forming classification functions that separated the 64 astronauts into their correct space sickness category with 77% success. The eight variables were chosen in this order: WETF training time, sitting systolic blood pressure, serum uric acid, calculated blood volume, serum phosphate, urine osmolality, environmental temperature at the launch site, and red cell count; serum chloride fell out from the earlier prediction. The jackknife cross-validation success was improved to 66%, and the subsample cross-validation to 71%.

For comparison, the percent success in predicting space sickness severity that would result from chance alone is 32%, as shown in Table 3. This table was computed by chi-square analysis, but the expected values in some cells are less than 1, violating the necessary assumptions for probabilistic comparisons with the above predictions of space sickness severity with and without WETF training. Combining the MODERATE and SEVERE groups to increase the expected values enabled valid chi-square comparisons for the classification and jackknife matrixes but not the subsample matrixes; for both those sets of predictions with and without WETF training, the predictions are significantly better than chance at p < 0.00005.

When WETF training time was used as the sole input variable for discriminant analysis to predict space sickness severity, the prediction successes obtained by the classification functions and by jackknife cross-validation were both 42%. The subsample cross-validation showed 33% success, the same expected due to chance.

Finally, the 64 astronauts were split into two groups according to whether they had WETF training (n = 36) or not (n = 28). Mann-Whitney tests between these two groups were performed on the nine fluid shift variables and 46 other clinical and anthropometric variables, some derived by calculation, available for this sample of astronauts. Of interest are these differences (mean ± SEM) found in serum chloride, WETF = 104.25 ± 0.41 mEq/L, no WETF = 105.86 ± 0.42 mEq/L

(p = 0.0153); plasma osmolality, WETF = 289.53 ± 0.62 mOsm/kg, no WETF = 292.21 ± 0.63 mOsm/kg (p = 0.0052); urine specific gravity, WETF = 1.0192 ± 0.0011, no WETF = 1.0162 ± 0.0012; and forced vital capacity indexed to body weight, WETF = 0.0660 ± 0.0018 L/kg (n = 35), no WETF = 0.0604 ± 0.0019 L/kg (n = 27), (p = .0400). These were the only variables found to differ between WETF trained and untrained astronauts at $p<0.05$.
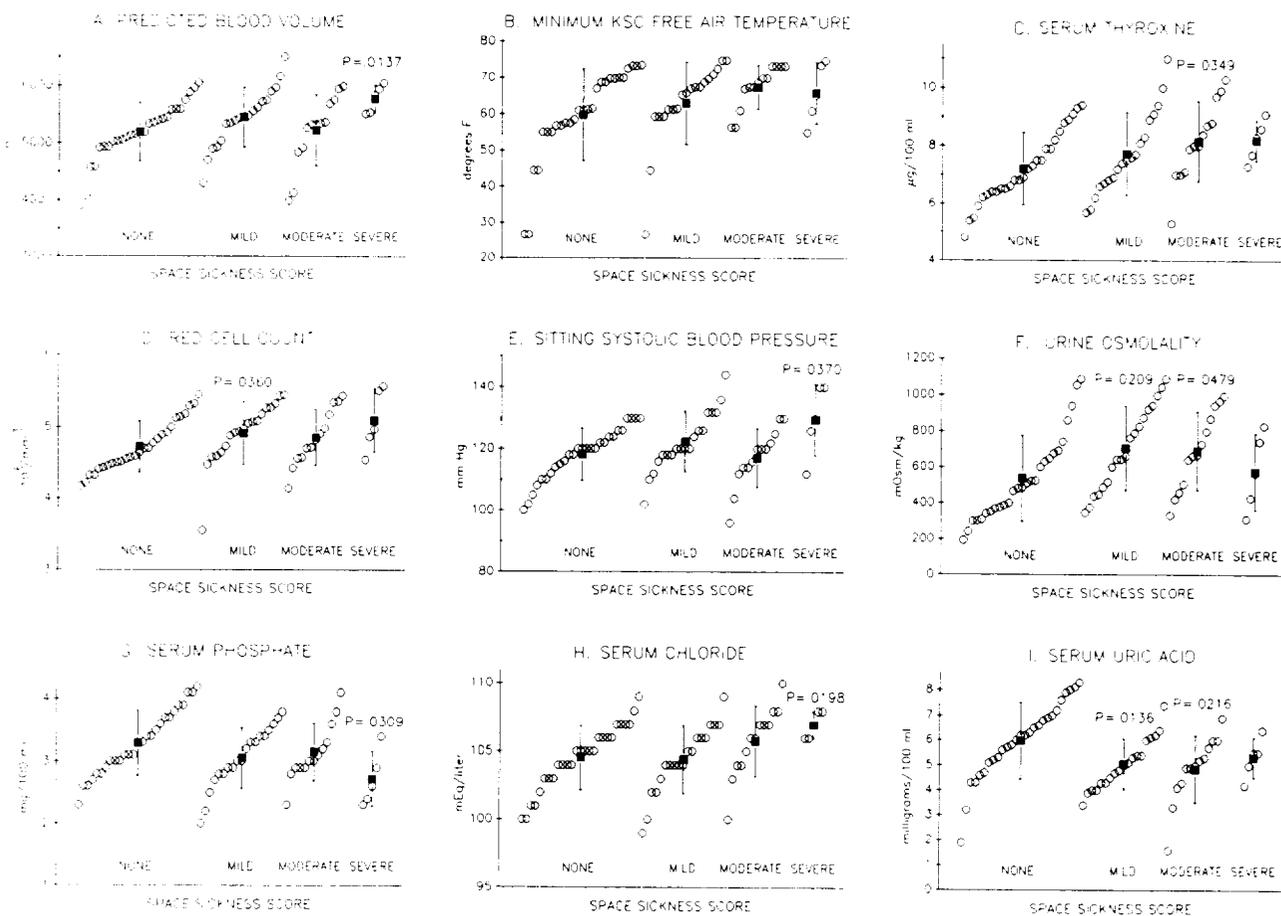

FIG. 1. NINE PREFLIGHT VARIABLES USED TO PREDICT SPACE SICKNESS.

Individual points are plotted in ascending order within each space sickness group at arbitrary but equidistant points along the x-axis for evaluation of the raw data. Black squares are means (± S.D.) for each group plotted at the median position in the range of points. P values are from Mann-Whitney significance tests of the three sick groups separately tested against the NONE group.
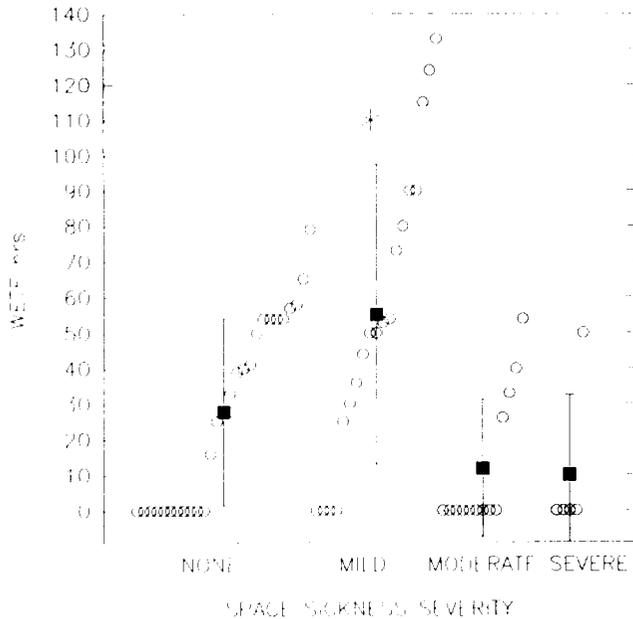
FIG. 2. HOURS IN WETF TRAINING VERSUS SPACE SICKNESS SEVERITY.

* The MILD group had significantly more training hours than all of the other groups by a Mann-Whitney test at p<0.05. Black squares are means (± S.D.).


FIG. 3. PRESENCE AND ABSENCE OF WETF TRAINING RELATIVE TO SPACE SICKNESS SEVERITY.

When the MODERATE and SEVERE groups were combined to enable a valid chi-square test, the differences in space sickness severity due to WETF training were significantly different from chance at p = 0.0067.

TABLE 2. SPEARMAN CORRELATIONS AMONG THE ORIGINAL NINE PREDICTION VARIABLES RELATED TO FLUID, ELECTROLYTE, OR CARDIOVASCULAR STATUS; INCLUDED ARE WETF TRAINING HOURS AND SPACE SICKNESS SCORES.

	SCORE	URICAC	THT4	CL	RBC	MINT	PO4	UROS	BV	SITSYS
URICAC	-.32*	1								
THT4	.32*	-.17	1							
CL	.27*	-.08	.29*	1						
RBC	.26*	.25*	.18	-.02	1					
MINT	.25*	-.07	.06	.14	-.20	1				
PO4	-.25*	-.05	.03	-.21	-.12	.10	1			
UROS	.24	.00	.17	.01	.21	.12	-.13	1		
BV	.23	.14	.20	-.05	.28*	.16	-.11	.27*	1	
SITSYS	.14	.17	-.09	-.14	.33*	-.11	-.10	.17	.11	1
WETF	-.17	.04	-.06	-.28*	.01	-.17	.05	.11	-.07	-.02

* = p < 0.05. All pairs are n = 64. SCORE = space sickness score; URICAC = serum uric acid; THT4 = serum thyroxine; CL = serum chloride; RBC = red cell count; MINT = minimum free air temperature at the launch site; PO4 = serum phosphate; UROS = urine osmolality; BV = predicted blood volume; SITSYS = sitting systolic blood pressure; WETF = hours of WETF training.

TABLE 3. PREDICTION SUCCESS FOR SPACE SICKNESS SEVERITY THAT WOULD BE EXPECTED SOLELY DUE TO CHANCE (determined by chi-square).

	PERCENT	PREDICTED			
CHANCE	CORRECT	NONE	MILD	MODERATE	SEVERE
NONE	47%	12.7	7.2	5.9	1.3
MILD	26%	8.9	5.0	4.2	0.9
MODERATE	21%	6.1	3.5	2.8	0.6
SEVERE	4%	2.3	1.3	1.1	0.2
TOTAL	32%	30	17	14	3

TABLE 4

PREDICTION OF SPACE SICKNESS SEVERITY USING NINE VARIABLES RELATED TO FLUID, ELECTROLYTE, AND
CARDIOVASCULAR STATUS SHOWN IN FIGURE 1

I. MEANS AND STANDARD DEVIATIONS OF VARIABLES USED

GROUP =	NONE	MILD	MODERATE	SEVERE
VARIABLE				
BV ml	5191 509	5449 524	5223 620	5770 242
MINT deg. F	59.8 12.6	63.0 11.3	67.5 6.0	65.9 8.4
THT4* µg/dL	7.20 1.25	7.73 1.43	8.16 1.38	8.20 .71
RBC 10^{12}/L	4.72 .36	4.91 .44	4.85 .39	5.09 .44
SITSYS mmHg	118 8	122 10	117 10	130 12
UROS mOsm/kg	536 241	706 234	692 218	574 215
PO4 mg/dL	3.29 .52	3.05 .49	3.15 .47	2.72 .44
CL mEq/L	105 2	104 3	106 3	107 1
URICAC mg/dL	5.97 1.53	5.06 1.02	4.86 1.35	5.32 .80
n	27	19	13	5

* THT4 was provided as input but not chosen by the
discriminant analysis program

II. CLASSIFICATION FUNCTIONS

GROUP =	NONE	MILD	MODERATE	SEVERE
VARIABLE				
BV	0.02517	0.02621	0.02502	0.02804
MINT	0.96085	1.05467	1.10696	1.10827
RBC	29.40623	32.22703	33.21688	33.39625
SITSYS	1.97799	2.03264	1.97129	2.15279
UROS	-0.01940	-0.01654	-0.01594	-0.02076
PO4	31.37756	29.48028	29.95746	28.24583
CL	21.55503	21.49715	21.71196	22.00787
URICAC	-2.07219	-3.45289	-3.40062	-3.82808
CONSTANT	-1448.37415	-1462.26501	-1481.95349	-1544.08777

III. CLASSIFICATION MATRIX

ACTUAL	PERCENT CORRECT	PREDICTED			
		NONE	MILD	MODERATE	SEVERE
NONE	85%	23	3	1	0
MILD	68%	4	13	2	0
MODERATE	39%	1	7	5	0
SEVERE	60%	0	1	1	3
TOTAL	69%	28	24	9	3

IV. JACKKNIFE CROSS-VALIDATION MATRIX

ACTUAL	PERCENT CORRECT	PREDICTED			
		NONE	MILD	MODERATE	SEVERE
NONE	81%	22	3	2	0
MILD	58%	5	11	3	0
MODERATE	23%	2	8	3	0
SEVERE	40%	0	2	1	2
TOTAL	59%	29	24	9	2

V. SUBSAMPLE CROSS VALIDATION ON 24 CASES NOT USED IN CALCULATION OF THE CLASSIFICATION FUNCTION

ACTUAL	PERCENT CORRECT	PREDICTED			
		NONE	MILD	MODERATE	SEVERE
NONE	100%	10	0	0	0
MILD	57%	2	4	0	1
MODERATE	20%	0	4	1	0
SEVERE	50%	0	1	0	1
TOTAL	67%	12	9	1	2

TABLE 5

PREDICTION OF SPACE SICKNESS SEVERITY USING NINE VARIABLES RELATED TO FLUID, ELECTROLYTE, AND CARDIOVASCULAR STATUS SHOWN IN FIGURE 1, WITH WETF TRAINING TIME ADDED

I. MEANS AND STANDARD DEVIATIONS OF VARIABLES USED

GROUP =	NONE	MILD	MODERATE	SEVERE
VARIABLE				
BV ml	5191 509	5449 524	5223 620	5770 242
MINT deg. F	59.8 12.6	63.0 11.3	67.5 6.0	65.9 8.4
THT4* μg/dL	7.20 1.25	7.73 1.43	8.16 1.38	8.20 .71
RBC 10^{12}/L	4.72 .36	4.91 .44	4.85 .39	5.09 .44
SITSYS mmHg	118 8	122 10	117 10	130 12
UROS mOsm/kg	536 241	706 234	692 218	574 215
PO4 mg/dL	3.29 .52	3.05 .49	3.15 .47	2.72 .44
CL* mEq/L	105 2	104 3	106 3	107 1
URICAC mg/dL	5.97 1.53	5.06 1.02	4.86 1.35	5.32 .80
WETF hrs	27.67 26.41	55.10 42.16	11.77 19.32	10.00 22.36
n	27	19	13	5

* THT4 and CL were provided as input but not chosen by the discriminant analysis program

II. CLASSIFICATION FUNCTIONS

GROUP =	NONE	MILD	MODERATE	SEVERE
VARIABLE				
BV	0.01556	1.11410	1.15981	1.16245
MINT	1.01556	1.11410	1.15981	1.16245
RBC	33.38202	36.10495	37.26220	37.48561
SITSYS	1.32721	1.39953	1.30835	1.48284
UROS	-0.00915	-0.00660	-0.00547	-0.01019
PO4	8.93290	7.10558	7.34489	5.32631
URICAC	-4.14167	-5.60553	-5.44387	-5.91047
WETF	0.08596	0.11994	0.07066	0.07598
CONSTANT	-222.64264	-246.32925	-237.16216	-265.42923

III. CLASSIFICATION MATRIX

ACTUAL	PERCENT CORRECT	PREDICTED NONE	MILD	MODERATE	SEVERE
NONE	89%	24	2	1	0
MILD	63%	5	12	2	0
MODERATE	77%	1	2	10	0
SEVERE	60%	0	1	1	3
TOTAL	77%	30	17	14	3

IV. JACKKNIFE CROSS-VALIDATION MATRIX

ACTUAL	PERCENT CORRECT	PREDICTED NONE	MILD	MODERATE	SEVERE
NONE	81%	22	2	3	0
MILD	58%	5	11	2	1
MODERATE	54%	3	3	7	0
SEVERE	40%	0	1	2	2
TOTAL	66%	30	17	14	3

V. SUBSAMPLE CROSS VALIDATION ON 24 CASES NOT USED IN CALCULATION OF THE CLASSIFICATION FUNCTION

ACTUAL	PERCENT CORRECT	PREDICTED NONE	MILD	MODERATE	SEVERE
NONE	100%	10	0	0	0
MILD	57%	2	4	0	1
MODERATE	40%	0	2	2	1
SEVERE	50%	0	1	0	1
TOTAL	71%	12	7	2	3

DISCUSSION

The validity of any prediction is limited by the reliability of the predictors and of the criterion variable, in this case space sickness score (Calkins et al., 1987). The possibility of errors in the predictor variables and the criterion should be recognized. Space sickness scores were evaluated by various flight surgeons over five years, based on medical debriefs from 64 individuals with varying subjective symptoms, many of whom were medicated in attempts to reduce the severity of their space sickness. It is possible that anti-motion sickness medication (primarily scopolamine-dextroamphetamine) may have reduced symptom severity in enough astronauts to bias the space sickness scores. However, because space sickness usually recurs in the same astronauts in subsequent flights (although there is a tendency for symptoms to lessen on a second flight), it is generally believed that space sickness scores are sufficiently reliable to be predictable (Calkins et al., 1987).

The reliability of the predictors probably varies from one to the next. Because the clinical variables were measured at different times before flight, few of them probably exactly matched the astronauts' physiologic status on launch day, with the possible exception of height, which was measured to the nearest inch, or half-inch in a few cases. Preflight blood volumes were only calculated, not measured. The free air temperature at the launch site is only a rough approximation of the actual environmental temperatures to which astronauts may have been exposed. Considering the potential sources of error in the data, it should not be surprising that even the best correlations of the predictor variables with space sickness are as low as they are. Conversely, detecting statistically significant relationships in the midst of noisy data implies the presence of fairly strong relationships.

Previous analyses (Simanonok et al., in review), if they hold up on further cross-validation, achieved NASA's technical goal of the development of a risk profile for predicting space sickness incidence with 80% confidence (Lackner, 1982). That rate of success in predicting space sickness severity, however, may await an understanding of the underlying mechanisms causing space sickness and improvements in the reliability of predictor and criterion scores, as has been proposed (Calkins et al., 1987).

Because the nine fluid shift variables describe aspects of preflight fluid, electrolyte, and cardiovascular status, applying them to predict space sickness with levels of success substantially better than chance is supportive of a fluid shift role in space sickness etiology. We cannot conclude if these predictor variables are markers for other factors or are somehow themselves determinants of space sickness, but some of them seem to present clearer relationships than others to a potential role of fluid shifts in space sickness, relationships which can be integrated into a mechanistic hypothesis for a fluid shift etiology of space sickness developed previously (Simanonok et al., in review). To summarize, the magnitude of an individual's response to fluid shifts probably depends on the magnitude of the fluid shift itself and on the individual's heart size relative to their blood volume. People with elevated blood volumes may suffer a greater central volume expansion in weightlessness, and vice versa. One response to central volume expansion on earth and in space is the release of atrial natriuretic peptides (ANP) from the heart (Epstein et al., 1987; Gharib et al., 1986; Leach et al., 1988). Mountain sickness may in some respects resemble space sickness because it is associated with central volume expansion and elevated plasma ANP (Bärtsch et al, 1991). And infusion of ANP at high doses into human subjects can cause nausea without any provocative motion stimulus at all (Weidmann et al., 1986). Therefore it is plausible that space sickness could result in part from a lowering of the threshold for nausea induction by ANP released in weightlessness in a dose-dependent fashion relative to individual responses to fluid shifts.

It is consistent with this fluid shift hypothesis of space sickness etiology that factors which modify the blood volume upward or downward might affect the volume of the fluid shift in weightlessness and subsequent physiologic responses. A variety of factors may modify the blood volume, some of which emerged previously among the nine predictor variables for space sickness (Simanonok et al, in review). In water immersion studies and head-down tilt as well as in weightlessness there is observed a contraction of the blood volume. It may be that with intermittent exposure to head-down tilt in WETF training, astronaut's blood volumes are reduced somewhat so that their subsequent fluid shifts in weightlessness are lessened in magnitude, hence their physiologic responses to fluid shifts are damped. WETF training may therefore be effective in ameliorating space sickness severity by partially preadapting the circulation to fluid shifts through a reduction of blood volume.

Alternatively or in addition, WETF training may be protective against space sickness because of the visual adaptation that it might provide. In WETF training, astronauts can assume almost any orientation with the spacecraft mockups. This could help to accustom them to the visual orientations of the spacecraft that they experience in space, thereby reducing their sensitivity to the "sensory conflict" that presently forms the dominant paradigm of space sickness etiology (Crampton, 1990).

The present data cannot be used to conclusively determine which of these two potential mechanisms of WETF training, preadaptation of the blood volume to fluid shifts or visual adaptation to weightless surroundings, plays the single or dominant role in ameliorating space sickness severity. However, the fact that the WETF trained group of astronauts had significantly lower serum chlorides and plasma osmolalities, and higher urine specific gravities and forced vital capacity indexes is consistent with probable effects of WETF training on fluid balance. There was also a low but significant inverse correlation between WETF training time and serum chloride concentration (Table 2). Lower serum chlorides and plasma osmolalities in WETF trained astronauts could indicate a state of partial recovery from fluid shifts, in which fluid and electrolyte losses caused by fluid shifts in WETF training were partially restored by drinking but electrolyte restoration lagged. Higher urine specific gravities in WETF trained astronauts would suggest renal states of fluid retention, increased electrolyte excretion, or thirst inhibition, any of

602

I

which might result from fluid shifts in WETF training depending on the timing of the training and the sampling. Forced vital capacity is greater in human subjects when standing compared to supine (Dikshit and Patrick, 1986) and after bed rest of 11 or 12 days (Beckett et al., 1986). This suggests that a decreased central blood volume allows greater lung capacities due to postural fluid shifts, after adaptation to fluid shifts in bedrest, and after adaptation to fluid shifts in WETF trained individuals.

Although mechanisms remain speculative, the data suggest that WETF training may be a partially effective method to reduce the severity of space sickness for astronauts. If WETF training is effective in ameliorating space sickness through visual adaptation, its role could be expanded so that all astronauts, or perhaps those who are predicted most likely to become space sick, get WETF training. If WETF training is shown to be beneficial as a countermeasure to space sickness by reducing the preflight blood volume and thereby diminishing the physiologic effects of fluid shifts in weightlessness, other more direct interventions to preadapt the circulation to fluid shifts may be even more effective.

REFERENCES

Bärtsch P, Pfluger N, Audétat M, Shaw S, Weidmann P, Vock P, Vetter W, Rennie D, Oelz O. Effects of slow ascent to 4559 M on fluid homeostasis. Aviat. Space Environ. Med. 1991; 62:105-10.

Beckett WS, Vroman NB, Nigro D., Thompson-Gorman S, Wilkerson JE, Fortney SM. Effect of prolonged bed rest on lung volume in normal individuals. J. Appl. Physiol. 1986; 61:919-25.

Calkins DS, Reschke MF, Kennedy RS, Dunlop WP. Reliability of provocative tests of motion sickness susceptibility. Aviat. Space Environ. Med. 1987; 58:A50-4.

Crampton GH, ed. Motion and Space Sickness. CRC Press, Boca Raton, 1990.

Davis JR, Vanderploeg JM, Santy PA, Jennings RT, Stewart DF. Space motion sickness during 24 flights of the Space Shuttle. Aviat. Space Environ. Med. 1988; 59:1185-9.

Dikshit MB, Patrick JM. Forced expiratory flow-volume curves during the application of lower-body negative pressure. Bull. Eur. Physiopathol. Respir. 1986; 22:599-603.

Epstein, M, Loutzenhiser R, Friedland E, Aceto R, Camargo MJF, Atlas S. Relationship of increased plasma atrial natriuretic factor and renal sodium handling during immersion-induced central hypervolemia in normal humans. J. Clin. Invest. 1987; 79:738-45.

Feldschuh J, Enson Y. Prediction of the normal blood volume: relation of blood volume to body habitus. Circulation 1977; 56:605-12.

Gharib C, Gauquelin G, Geelin G, Vincent M, Chaemmaghami F, Grange C, Cantin M, Gutkovska J, Guell A. Levels of atrial natriuretic factor (alpha h ANF) during acute simulated weightlessness. In: Hunt J (ed). Proc. of 2nd Intl. Conf. on Space Physiology, Toulouse, France, 20-22 Nov. 1985, European Space Agency publication no. SP-237, Paris, 1986, pp 173-6.

Greenleaf J. Physiology of fluid and electrolyte responses during inactivity: water immersion and bed rest. Med. Sci. Sports Exerc. 1984; 16:20-5.

Homick JL, Reschke MF, Vanderploeg JM. Prediction of susceptibility to space motion sickness. In Graham MD and JL Kemink. eds. The Vestibular System: Neurophysiologic and Clinical Research, Raven Press, New York, 1987:39-49.

Kakurin LI, Kuzmin MP, Matsnev EI, Mikhailov VM. Physiological effects induced by antiorthostatic hypokinesia. In: Life Sciences and Space Research XIV; Proceedings of the Open Meeting of the Working Group on Space Biology, May 29-June 7, 1975, and Symposium on Gravitational Physiology, Varna, Bulgaria, May 30, 31, 1975. Akademie-Verlag GmbH, Berlin, East Germany, 1976:101-8

Lackner JR. Comments on prediction subgroup meeting. In: Homick JL, ed. Space Motion Sickness Workshop Proceedings June 21-22, 1982. NASA Johnson Space Center publication JSC 18681, 1982:5-11, 76.

Lathers CM, Charles JB, Elton KF, Holt TA, Mukai C, Bennett BS, Bungo MW. Acute hemodynamic responses to weightlessness in humans. J. Clin. Pharmacol. 1989; 29:615-27.

Leach CS, Johnson PC, Cintron NM. The endocrine system in space flight. Acta Astronautica 1988; 17:161-6.

Mitarai G, Mano T, Yamazaki Y. Correlation between vestibular sensitization and leg muscle relaxation under weightlessness simulated by water immersion. Acta Astronautica 1981; 8:461-8.

Reschke MF, Homick JL, Ryan P, Moseley EC. Prediction of the space adapation syndrome. Proceedings of NATO Conference No. 372 of the Advisory Group for Aerospace Research and Development. Motion sickness: mechanisms, prediction, prevention and treatment, 1984:26-1 to 27-19.

Simanonok KS, Moseley EC, Davis JR, Charles JB. Space sickness predictors and the fluid shift. In review.

Thornton WE, Moore TP, Pool SL, Vanderploeg J. Clinical characterization and etiology of space motion sickness. Aviat. Space Environ. Med. 1987; 58:A1-8.

Vanderploeg JM, Stewart DF, Davis JR. Space motion sickness. In: Hunt J (ed). Proc. of 2nd Intl. Conf. on Space Physiology, Toulouse, France, 20-22 Nov. 1985, European Space Agency publication no. SP-237, Paris, 1986:137-42.

Weidmann P, Hasler L, Gnadinger MP, Lang RE, Uehlinger DE, Shaw S, Rescher W, Reubi FC. Blood levels and renal effects of atrial natriuretic peptide in normal man. J. Clin. Invest. 1986; 77:734-42.

Youmans EM, Charles JB, Santy PA. The relationship between preflight underwater training and space motion sickness (abstract). Aviat. Space Environ. Med. 1987; 58:497.

COMPUTER SIMULATION OF PREFLIGHT BLOOD VOLUME REDUCTION
AS A COUNTERMEASURE TO FLUID SHIFTS IN SPACEFLIGHT

K.E. Simanonok, R. Srinivasan, and J.B. Charles

NASA/Johnson Space Center, Houston, TX 77058

ABSTRACT

Fluid shifts in weightlessness may cause a central volume expansion, activating reflexes to reduce the blood volume. Computer simulation was used to test the hypothesis that preadaptation of the blood volume prior to exposure to weightlessness could counteract the central volume expansion due to fluid shifts and thereby attenuate the circulatory and renal responses resulting in large losses of fluid from body water compartments. We used the Guyton Model of Fluid, Electrolyte, and Circulatory Regulation, modified to simulate the six-degree head-down tilt that is frequently used as an experimental analog of weightlessness in bedrest studies. Simulation results show that preadaptation of the blood volume by a procedure resembling a blood donation immediately before head down bedrest is beneficial in damping the physiologic responses to fluid shifts and reducing body fluid losses. After ten hours of head down tilt, blood volume after preadaptation is higher than control for 20 to 30 days of bedrest. Preadaptation also produces potentially beneficial higher extracellular volume and total body water for 20 to 30 days of bedrest.

INTRODUCTION

The mechanisms causing an increased excretion of water and electrolytes in ground-based experiments simulating weightlessness by water immersion or head-down tilt are thought to be initiated by the headward redistribution of body fluids (Leach, 1979; Epstein et al., 1980). In weightlessness, circulatory changes include an early and rapid inflight elevation of hemoglobin concentration (Kimzey, 1977) coinciding with a loss of body mass that reflects a negative water balance. Most of the water loss occurs within the first three days, and a plateau is thought to be reached by 10 or 20 days of weightlessness (Thornton and Ord, 1977). Red cell mass may decrease by 15% after 2 or 3 weeks in space (Nicogossian and Parker, 1982).

The reduction of blood volume in weightlessness is considered to be an adaptation to the space environment (Cogoli, 1981; Leach et al., 1983; Ilyushin and Burkovskaya, 1980). Unfortunately this leaves astronauts maladapted to gravity when they return to earth; postflight orthostatic intolerance has been a problem that is now par-

tially counteracted by oral rehydration before reentry (Bungo, 1989). It would be desirable to limit inflight body fluid losses if possible. There is also a possibility that fluid shifts play a role in space sickness etiology (Barrett and Lokhandwala, 1981; Leach, 1987; Kohl, 1985, 1987). Reducing the physiologic effects of fluid shifts might therefore help to ameliorate the space sickness problem.

The present study was undertaken to test the hypothesis that some of the major physiologic effects of a fluid shift could be counteracted by reducing the blood volume prior to the fluid shift. Blood volume reduction should decrease the central volume expansion due to fluid shifts and attenuate the circulatory, hormonal, and renal responses experimentally observed in water immersion and head-down tilt that are often used to model weightlessness. The net effects should be an attenuation of the physiologic responses to fluid shifts and decreased body fluid losses.

This hypothesis has previously been supported in two short-term water immersion experiments in human subjects (Simanonok and Bernauer, 1985; Simanonok, in review). For the present study we used a validated mathematical model of circulation to test the above hypothesis. Our purposes in employing mathematical modeling are 1) to extend the time period for testing the preflight blood volume reduction countermeasure beyond that possible in water immersion experiments, 2) to efficiently test different methods to optimize the countermeasure, and 3) to assist in the experimental design of further tests of the countermeasure using human subjects. This paper presents the initial results of our work toward those goals.

METHODS

The mathematical model used in this study was derived from the Guyton Model of Fluid, Electrolyte, and Circulatory Regulation (1972). It incorporates known relationships between physical, neural, and hormonal regulators of fluid balance and volume, pressure, and flow within the human circulatory system and fluid compartments (Fig. 1). The original model was modified for weightless simulation by head-down tilt by White (1974) and later improved for this purpose by Leonard and Grounds (1977). The model has been extensively validated through comparison with data

from ground-based and flight experiments (Leonard et al., 1979, 1986). The present version of the model is coded in FORTRAN and runs on the IBM PC and compatibles. It required two changes for this simulation.

The first modification of the model was to enable it to reach a stable equilibrium at a reasonable reduced blood volume with a normal hematocrit after prolonged head-down tilt. The rationale for this change was based on the assumption that there is a final adapted state in weightlessness in which a new fluid equilibrium exists; after adaptation, the blood volume and other body fluid compartments should no longer continue to decrease. Present evidence suggests that the magnitudes of these adaptive changes in true weightlessness are somewhat greater than in bedrest (Leach and Johnson, 1984), but the model changes posture to simulate bedrest, not weightlessness. We used the experimental data from a bedrest study (Leach and Johnson, 1984) which was specifically designed to simulate the 10 day Spacelab 1 mission. The red cell mass reduction after 10 days of bedrest was -3.3% ± SEM 1.6% (n=6) as compared to the losses in the four Spacelab 1 astronauts of -9.3% ± SEM 1.6%. Obtaining a stable equilibrium with the model of a reduced blood volume including a red cell mass loss of -3.3% at 10 days required the replacement of the erythropoiesis block of the Guyton Model with a model of erythropoiesis developed by Leonard et al. (1981). The difference between the starting blood volume and the new blood volume at equilibrium was the preadaptation volume, the volume to remove to counteract fluid shifts.

The second change necessary was to enable simulation of bleeding. This required the addition of a BLEED RATE parameter allowing blood to be removed at any desired rate, with the removal of the constituents of blood (water, cells, salts, and protein) in proper proportions. From a previous test of the countermeasure in a water immersion study we had serial hematocrit data taken every two minutes during hemorrhages of 15% of blood volume in 8 supine subjects (unpublished data, Simanonok); two of these subjects had the same hemorrhage times of 18 minutes. We compared the data from these two subjects with a simulated bleed of 15% of blood volume in 18 minutes to validate the model's response to acute blood volume reduction.

To test our hypothesis that blood volume reduction would counteract fluid shifts, we simulated an acute blood volume reduction by the preadaptation volume over 18 minutes in the supine position immediately before a fluid shift induced by 6 degree head-down tilt. The control simulation was 6 degree head-down tilt alone.

RESULTS

The starting volume of blood before simulated head-down tilt was 5024 ml. After prolonged 6 degree head-down tilt (70+ simulated days), the equilibrium blood volume was 4490 ml. The difference, 534 ml, or about 11% of the starting blood volume, was the preadaptation volume to remove before head-down tilt to test the countermeasure.
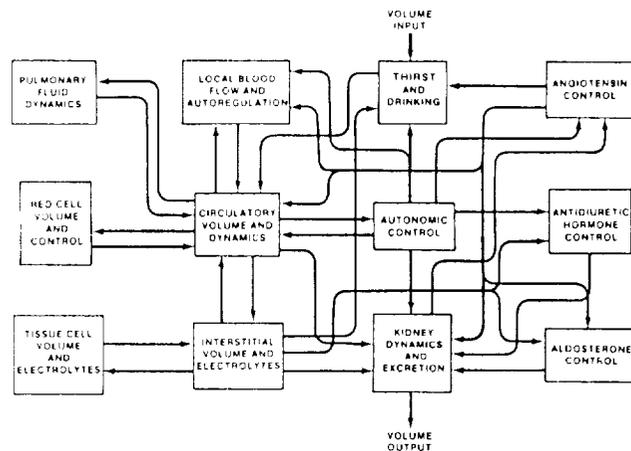

Fig. 1.   Subsystems of the Guyton Model of Circulatory, Fluid, and Electrolyte Regulation.

The protective effects of preadapting the circulation to fluid shifts in maintaining blood volume, extracellular fluid volume, and total body water are shown in Figs. 2, 3, and 4, respectively. Of particular interest in Fig. 2 is that after ten hours of simulated head down tilt, the blood volume can be slightly better maintained for 20 to 30 days by a prior bleeding. The effects on urine flow are shown in Fig. 5. Data are plotted on a logarithmic time scale so that all phases of the experiment out to 70 days may be clearly distinguished.
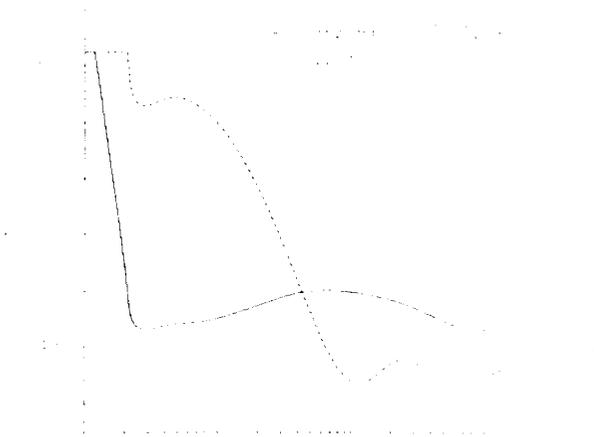

Fig. 2.   BLOOD VOLUME. BV REDUCTION = acute reduction of blood volume by 534 ml in 18 minutes. HDT = head down tilt.

606
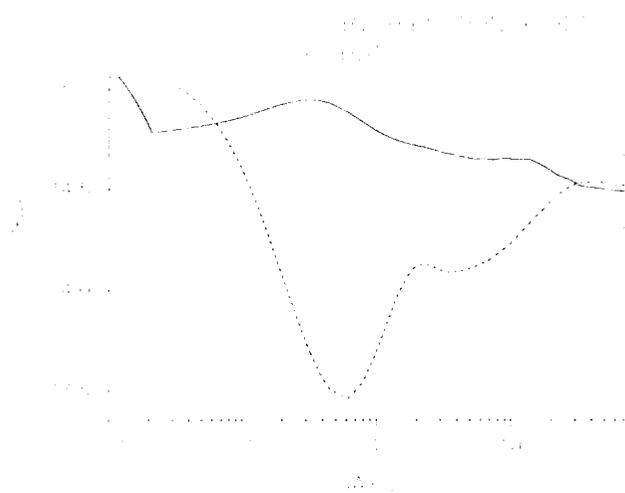
Fig. 3.   EXTRACELLULAR FLUID VOLUME
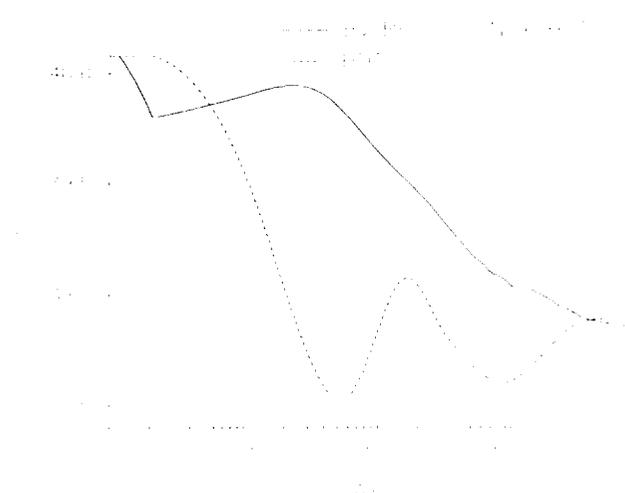

Fig. 4.   TOTAL BODY WATER
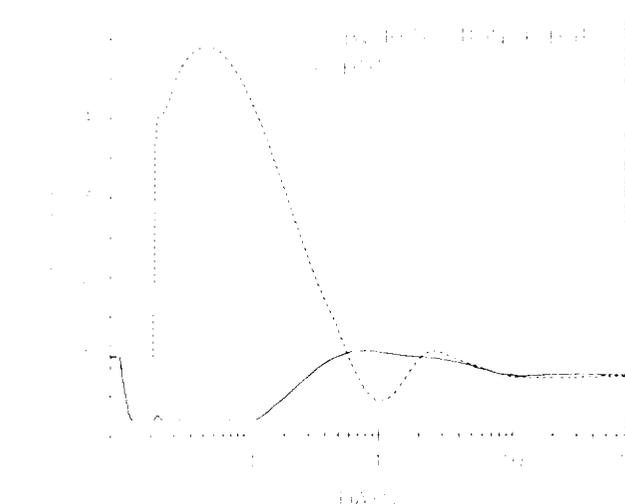
Fig. 5.   URINE FLOW

DISCUSSION

The hypothesis that blood volume reduction before a headward fluid shift would counteract some of the major effects of that fluid shift was supported by the results of this computer simulation. These results are in agreement with two earlier water immersion studies (Simanonok and Bernauer, 1985; Simanonok, in review). Except for a short period of time after blood volume reduction during the first few hours of head down tilt, this simulation suggests that body fluid volumes would be better maintained for twenty to thirty days by preadapting the blood volume to headward fluid shifts. During the first few hours, loss of fluid by elevated urine flow (Fig. 5) in head-down tilt without the countermeasure rapidly met and surpassed the initial fluid loss due to a prior blood volume reduction.

In addition to attenuation of the volume losses, Figs. 2, 3, and 4 illustrate that the rates of change of the body fluid volumes during head down tilt are also decreased by prior blood volume reduction. Decreases in both the rates of change and the absolute magnitudes of the changes in body fluid volumes suggest that the physiologic stresses upon the simulated body were decreased. Homeostasis is the term applied to the physiologic tendency to maintain a constancy of the body's internal environment. Fluid shifts disturb the internal environment, invoking homeostatic responses that act to return the parameters of the internal environment to their normal operating range. Preadaptation of the blood volume may therefore reduce the magnitude of the homeostatic stresses accompanying adaptation to weightlessness.

The most unambiguous potential benefit of the countermeasure is that postflight orthostatic tolerance might be improved by the enhanced fluid retention for missions up to twenty or thirty days in length. There may also be other benefits. Head congestion, facial edema, and headaches in astronauts might be lessened. More important, if it is true that fluid shifts in some manner contribute to space sickness etiology (Barrett and Lokhandwala, 1981; Leach, 1987; Kohl, 1985, 1987), then it is possible that preadaptation of the circulation may be beneficial in ameliorating space sickness by reducing the physiologic impact of fluid shifts. Unfortunately, the mechanisms causing space sickness are only speculative and there are no nausea or motion sickness variables in the Guyton Model.

It is important to consider other ramifications of blood volume reduction that may be relevant to astronauts. We are not proposing that the optimum method to preadapt the circulation to weightlessness is by acute hemorrhage immediately preflight, in a procedure resembling a blood donation. This is because blood volume is usually restored within a day or two after an acute hemorrhage, although the restoration of plasma proteins and cells takes longer. Launch delays could easily ruin the effectiveness of that approach. A preflight regimen of diuretics, water immersion, or head-down tilt might effectively reduce the blood volume; at least one astronaut has slept in head-down tilt for about ten days before each of his three Shuttle missions to reduce his blood volume and ameliorate the effects of fluid shifts. He

607

reports that he suffered none of the head congestion, facial edema, or headaches that astronauts commonly experience in weightlessness, and he never became space sick (Musgrave, personal communication).

Preflight orthostatic tolerance would if anything be decreased by blood volume reduction, so astronauts who were borderline in their orthostatic tolerance should probably not be preadapted unless their orthostatic tolerance could first be improved. Exercise capacity could be reduced somewhat by preflight blood volume reduction, and could conceivably affect an astronaut's performance in the event of a preflight emergency egress. These concerns must be thoroughly addressed before preflight blood volume reduction can be seriously considered for astronauts. Further computer simulations may be useful in answering some of these questions.

The utility of computer simulation in planning and designing experimental studies is amply demonstrated by this and previous simulation studies. Simulation is a highly cost-effective method for initial hypothesis testing and experimental design optimization, especially when investigating the physiology of human adaptation to weightlessness.

This computer simulation extends earlier work which suggests that preflight blood volume reduction on the order of magnitude of an ordinary blood donation could be beneficial to astronauts by damping their physiologic responses to fluid shifts and reducing body fluid losses. These results suggest that this countermeasure to fluid shifts could possibly improve postflight orthostatic tolerance for missions of 20 to 30 days in duration, and could ameliorate some of the head congestion, facial edema, and headaches that afflict many astronauts. To the extent that fluid shifts contribute to the problem, an effective countermeasure to fluid shifts could also be beneficial against space sickness.

REFERENCES

Barrett RJ, Lokhandwala MF. Circulatory and vestibular implications of central angiotensin mechanisms in physiological adaptation to weightlessness. Medical Hypotheses 1981; 7:1415-9.

Bungo MW. The cardiopulmonary system. In: Nicogossian AE, Huntoon CL, Pool SL, eds. Space Physiology and Medicine, 2nd ed. Lea & Febiger, Philadelphia, 1989:179-201.

Cogoli A. Hematological and immunological changes during space flight. Acta Astronaut. 1981; 8:995-1002.

Epstein M, DeNunzio AG, M Ramachandran. Characterization of renal response to prolonged immersion in normal man. J. Appl. Physiol. 1980; 49:184-8.

Guyton AC, Coleman TG, Granger HJ. Circulation: overall regulation. Ann. Rev. Physiol. 1972; 34:13.

Ilyushin AV, Burkovskaya T Ye. Cytokinetic evaluation of erythropoiesis during long-term orbital flights. USSR Report: Space Biology and Aerospace Medicine 1981; 15:60-4.

Kimzey SL. Hematology and immunology studies. In: Johnson RS, Dietlin LF, eds. Biomedical Results from Skylab. NASA SP-377, Sup. of Documents, Washington, DC, 1977:249-82.

Kohl RL. Endocrine correlates of susceptibility to motion sickness. Aviat. Space Environ. Med. 1985; 56:1158-65.

Kohl RL. Hormonal responses of metoclopramide-treated subjects experiencing nausea or emesis during parabolic flight. Aviat. Space Environ. Med. 1987; 58:A266-9.

Leach CS. A review of the consequences of fluid and electrolyte shifts in weightlessness. Acta Astronaut. 1979; 6:1123-35.

Leach CS. Fluid control mechanisms in weightlessness. Aviat. Space Environ. Med. 1987; 58:A74-9.

Leach CS, Johnson PC. Influence of spaceflight on erythrokinetics in man. Science 1984; 225:216-8.

Leach CS, Johnson PC, Suki WN. Current concepts of space flight induced changes in hormonal control of fluid and electrolyte metabolism. The Physiologist 1983; 26:S24-7.

Leonard JI, Grounds DJ. Modification of the Long Term Circulatory Model for the Simulation of Bed Rest. NASA Contractor Report (NASA CR-160186), Washington, D.C., 1977.

Leonard JI, Kimzey SL, Dunn CDR. Dynamic regulation or erythropoesis: A computer model of general applicability. Exp. Hematol. 1981; 9:355-78.

Leonard JI, Rummel JA, Leach CS. Computer simulations of postural change, water immersion and bed rest: an integrated approach for understanding the spaceflight response. The Physiologist, 22:S32-2, 1979.

Leonard JI, White RJ, Rummel JA. Math modelling as a complement to the scientific inquiry of physiological adaptation to space flight: fluid, endocrine and circulatory regulation. In: Hunt J, ed. Proc. of 2nd Intl. Conf. on Space Physiology, Toulouse, France, 20-22 Nov. 1985, European Space Agency publication no. SP-237, Paris, 1986:233-44.

Nicogossian AE, Parker JF Jr. Space Physiology and Medicine. NASA SP-447, Sup. of Documents, Washington, D.C., 1982:129.

Simanonok KE. Effects of hypovolemia on the responses to water immersion in men. In review.

Simanonok KE. Bernauer EM. Hemorrhage as a countermeasure to fluid shifts in water immersion. (abstract) The Physiologist 1985; 28:285.

Thornton WE, Ord J. Physiological mass measurements in Skylab. In: Johnson RS, Dietlin LF, eds. Biomedical Results from Skylab. NASA SP-377, Sup. of Documents, Washington, D.C., 1977:175-90.

White RJ. A Long-term Model Of Circulation: Final Report. NASA Contractor Report (NASA CR-147674), Washington, D.C., 1974.

Abstract for SOAR '91 Conference

1990 Hypobaric Decompression Sickness Workshop:
Summary and Conclusions

Andrew A. Pilmanis Ph.D*, Barbara J. Stegmann MD**,
and Terrell E. Scoggins*

* High Altitude Protection Function
USAF Armstrong Laboratory/VNBD
Brooks AFB TX 78235-5000

** KRUG Life Sciences
P.O. Box 790644
San Antonio TX 78279

Decompression sickness resulting from exposure to the
hypobaric environment was reviewed and discussed at a three-day
workshop hosted by the US Air Force Armstrong Laboratory in
October 1990 with co-sponsorship from NASA Johnson Space Center
and the Air Force Office of Scientific Research. This milestone
meeting, attended by over 50 participants representing the
Department of Defense, NASA, European Space Agency and academia,
updated the current understanding of altitude decompression
sickness (DCS). Both research and operational aspects of this
illness were addressed through presentations on the pathophysiol-
ogy and clinical manifestations of DCS, its incidence in aviation
and space operations, and existing and proposed measures for DCS
prevention. Specific areas requiring further research were also
identified. This paper summarizes the material presented at this
workshop.

Note: This paper will also be presented at the AGARD 71st
Aerospace Medical Panel Symposium; 29 Apr-3 May 1991, Pensacola FL
and published in the symposium procedings.

Improving Survival After Tissue Vaporization (Ebullism).

Barbara J. Stegmann, M.D., KRUG Life Sciences, San Antonio, TX
Andrew A. Pilmanis, Ph.D., Armstrong Laboratories, Brooks AFB, TX
Toniann Derion, Ph.D., NRC Resident Research Associate, Brooks AFB, TX

Exposure of unprotected humans to altitudes above 63,000 feet results in ebullism. Ebullism occurs when the vapor pressure of tissues is less than the ambient pressure and the tissues spontaneously "boil". This condition results in rapid loss of consciousness, cardiac "vaporlock", pulmonary collapse, cerebral anoxia and, in the absence of recompression, death. Potential scenarios for ebullism include extravehicular activity (EVA) accidents in space, aircraft experiencing rapid decompressions at high altitudes with cabin or pressure suit failure, and accidents during pressure suit training exercises in altitude chambers such as those exercises required for high altitude reconnaissance qualifications. However, there are no established medical protocols for the treatment of this condition. As the injury pattern of ebullism is better defined, improved protective measures and in-depth treatment protocols will be developed.

The pathophysiology of ebullism was studied from the 1940s through the 1960s using animal models. Survival depended upon exposure times at altitude (30 seconds to 2.5 minutes) and was species specific. For example, chimpanzees survived 2.5-minute exposures to 120,000 feet and showed no drop in baseline neurologic task performance immediately after the exposure. However, rats did not survive exposures greater than 40 seconds.

There is one report in the literature of a prolonged, unprotected, human exposure and several anecdotal reports of unprotected, short-term human exposures to near-vacuum. During these exposures, the individuals rapidly lost consciousness; however, the degree of sustained injury ranged from insignificant barotrauma to massive cerebral and pulmonary injury requiring intensive medical intervention. Reports of experiments in which subjects had partial protection against near-vacuum described swelling and pain in the unprotected limbs. These individuals required minor medical care and no resuscitation.

In addition to pulmonary and neurologic concerns, unprotected exposure of the head may result in freezing of the corneal surface of the eye. Surface eye freezing may impair vision and significantly impact mission completion. The type and amount of damage done to the eye by this type of freezing is not known.

At this time, little data are available on the effectiveness of conventional treatment protocols, such as hyperbaric oxygen (HBO), for ebullism-induced injuries. The Armstrong Laboratory at Brooks Air Force Base is currently evaluating the use of HBO in ebullism. Also, research is needed to assess the efficacy of other adjunctive therapies such as high frequency ventilation and cerebral protective drugs that are still under development. Survival rates are expected to improve as a more complete understanding of the pathophysiology of ebullism is attained, as development of new protective measures progresses, and with the development of specific medical protocols.

Session H7: SPACE CABIN CONTAMINANTS

Session Chair: Dr. Jeffrey W. Fisher

TOXICOLOGICAL APPROACH TO SETTING SPACECRAFT MAXIMUM ALLOWABLE CONCENTRATIONS FOR CARBON MONOXIDE

K.L. Wong[*], T.F. Limero[*], and J.T. James.
NASA-JSC Biomedical Operations and Research Branch,
Johnson Space Center, Houston, Texas.
[*]Krug Life Sciences, Houston, Texas.

ABSTRACT

The Spacecraft Maximum Allowable Concentrations (SMACs) are exposure limits for airborne chemicals used by the National Space and Aeronautics Administration (NASA) in spacecraft. The aim of these SMACs is to protect the spacecraft crew against adverse health effects and performanace decrements that would interfere with mission objectives. Because the 1-hr and 24-hr SMACs are set for contingencies, minor, reversible toxic effects that do not affect mission objectives are acceptable. The 7-day, 30-day, or 180-day SMACs are aimed at nominal operations, so they are established at levels that would not cause noncarcinogenic toxic effects and more than one case of tumor per 1000 exposed individuals over the background. The process used to set the SMACs for carbon monoxide (CO) is described to illustrate the approach used by NASA.

After the toxicological literature on CO was reviewed, the data were summarized and separated into acute, subchronic, and chronic toxicity data. CO's toxicity depends on the formation of carboxyhemoglobin (COHb) in the blood, reducing the blood's oxygen-carrying capacity. The initial task was to estimate the COHb levels that would not produce toxic effects in the brain and heart. Then the Coburn-Forster-Kane (CFK) equation was used to calculate the CO exposure concentration that would lead to a no-effect COHb level upon exposure for a given duration (1 hour to 180 days) and that CO concentration was selected as the SMAC. The 1-hr and 24-hr SMACs were set at 55 ppm and 20 ppm, respectively, to prevent central nervous system (CNS) impairment at COHb of 3%. Ten ppm was selected to be the 7-day, 30-day, and 180-day SMACs to ensure that the COHb level would not exceed 1.6%, which would protect against abnormal heart rhythms and CNS impairment.

1. INTRODUCTION

Inhalation exposures to some chemicals may lead to toxic effects that impair the health or performance of the exposed subjects. To provide a safe environment for the astronauts in spacecraft, NASA has established SMACs for 190 chemicals[1]. The SMACs are exposure limits that, if not exceeded, would prevent health injuries and the crew's performance decrement due to exposures to airborne chemicals. The official SMACs were set, without documentation, about 20 years ago and currently they are designated as 7-day SMACs for space shuttle missions.

There are two reasons why the official 7-day SMACs need to be revised: 1) availability of new toxicity data for most of these chemicals and 2) scientists have a better understanding of toxicologic principles. In addition to revision of the 7-day SMACs, a range of SMACs is necessary to a adequately cover all potential exposures associated with nominal operations, contingency events and extended duration missions. The Johnson Space Center (JSC) Toxicology Group has decided that this range should include SMACs for 1 hr, 24 hr, 30 days, and 180 days. The 7-day, 30-day, and 180-day SMACs are designed for nominal operations, so they are set at levels that would not cause noncarcinogenic toxic effects. For carcinogens, the 7-day, 30-day, and 180-day SMACs are established so that the risk of getting tumors is no more than one in a thousand above the background level. In contrast, the 1-hr and 24-hr SMACs are aimed at contingency situations. Therefore, minor, reversible toxic effects are acceptable provided that they do not prevent the spacecraft crew from achieving the mission objectives. Examples of these acceptable toxic effects are mild mucosal irritation and headaches. Compared with the longer-term SMACs, the 1-hr and 24-hr SMACs are, therefore, sometimes set at levels relatively higher than what the differences in exposure duration would suggest.

The JSC Toxicology Group has embarked on the project of setting new 1-hr, 24-hr, 7-day, 30-day, and 180-day SMACs for the 190 chemicals in the current official list and a number of new chemicals. The new SMACs will be used in support of the Space Shuttle missions and future Space Station Freedom missions. In addition, some of the new SMACs will serve as the design criteria for the air revitalization subsystem of the Environment Control and Life Support System and for air-quality monitoring equipment of the Environmental Health System in the Space Station Freedom. Overly conservative SMACs must be avoided because these could cause these systems to be designed with unnecessary weight or power

burden.

Establishing new SMACs proceeds in three stages. First, the JSC Toxicology Group sets interim SMACs based on its understanding of the chemical toxicity and state-of-the-art toxicologic principles, assisted by a guideline provided by the National Research Council's Committee on Toxicology (NRC-COT). The interim SMACs and their rationale are, then, reviewed by the NRC-COT. Finally, the JSC Toxicology Group would, if necessary, revise the interim SMAC values and the documentation according to the NRC-COT's comments. The interim SMACs become official after this review process.

This paper describes the JSC Toxicology Group's approach in setting the new SMACs by using carbon monoxide as an example. CO is a colorless and odorless gas[2]. It is produced both endogenously and exogenously. In the body, CO is produced at a rate of 0.4 ml/hr via catabolism of hemoglobin and nonhemoglobin heme[3]. The incomplete oxidation during thermodegradation of materials containing carbon in an atmosphere containing oxygen is an exogenous source of CO. For instance, CO could be produced when gas stoves and furnaces are in use. Smoking is another common indoor exogenous source of CO[4], which has been measured in main-stream cigarette smoke at 40,000-50,000 ppm[5]. The primary outdoor source of CO is automobile exhaust[4]. In the 1960's and 1970's, CO concentrations of 20-30 ppm were commonly detected on expressways in major U.S. cities and on a busy street inside London during rush hours, with peak concentrations exceeding 100-300 ppm[4,5].

Inside spacecraft, there is no known use of carbon monoxide. However, CO metabolically produced by the astronauts may accumulate in the cabin. Of course, a large amount of carbon monoxide may be produced if there were a thermodegradation accident in the cabin. Therefore, carbon monoxide could be a threat to the crew's health and SMACs for carbon monoxide are critically needed for crew protection.

2. GENERAL APPROACH IN SETTING SMACs

2.1. Literature Searches

2.1.1. Nonbiological Literature

The SMAC-setting process begins with a literature search on the important physical and chemical properties of the chemical. Information gathered typically includes the Chemical Abstract registry number, chemical structure, molecular weight, boiling point, melting point, and vapor pressure. The Chemical Abstract registry number is to facilitate the toxicological literature search. The chemical structure sometimes gives a preliminary idea of the chemical's toxicity and it, thus, also helps to plan the toxicological literature search. The molecular weight is used in the conversion of airborne concentration units from mg/m^3 to ppm or vice versa.

The boiling and melting points determine the potential physical form of the chemical in the spacecraft. The physical form, in turn, determines the exposure pathway. For instance, if the chemical exists as a solid, the potential exposure routes are ingestion, cutaneous contact, and inhalation as a dust. Since ingestion can be avoided if care is exercised by the astronauts and cutaneous contacts with a solid usually will not present a systemic toxic hazard, the most significant exposure route of a solid is by inhalation, provided that the solid exists as particulate. On the other hand, if the chemical is a liquid at room temperature and normal ambient pressure, it may be inhaled as a vapor, plus there may be cutaneous exposures. Of course, if the chemical exists as a gas, inhalation would be the primary route of exposure.

The vapor pressure gives an indication of the potential for the chemical to be inhaled. For example, if the vapor pressure is low, the liquid may have to be aerosolized before it would present a significant inhalation hazard. However, if the liquid is highly volatile, inhalation in the vapor form would be the most likely scenario. Although solid or liquid chemicals will float in air in microgravity, presenting a potential to be ingested, the chance of such an event occurring to a significant extend is very slim. Consequently, oral exposures are not considered in setting SMACs.

2.1.2. Biological Literature

The second step is to thoroughly search the literature on the toxicological properties of the chemical, using the computerized data bases of the National Library of Medicine. The particular data bases routinely searched are the TOXLINE, TOXLIT, IRIS, and the Hazardous Substances Data Bank. Occasionally, Chemical Abstracts is also searched. After an initial screen of the publication titles, publications of interest are selected and their abstracts are then extracted from the data bases. These abstracts are read and the key publications are obtained via our library.

2.1.3. Literature Selection Criteria

The selection criteria for publications include the type of study system used by the investigators, the species of the study subjects, and the exposure route used in the study. These criteria all could influence the usefulness of the publication in setting SMACs.

2.1.3.1. Study System as a Selection Criterion

Except for the studies of the effect of the chemical on genetic materials, the literature on the chemical's effect on an in vitro system is routinely not reviewed. In vitro systems are systems in which the experiments are conducted with tissues or cells isolated from the body and maintained in a growth media in a petri dish or flask. The reason in vitro studies are not relied upon is three-fold. First, not all the cell types are present in an in vitro system, so the response seen in the system may be different from the response seen in the body. Second, when

the tissues or cells are maintained in a growth media, the tissues or cells may change their biological characteristics so that they react to the chemical differently than they would have in the body. Third, it is difficult to convert the chemical concentration used in the growth media to an inhalation exposure concentration. Due to the artificial nature and all the uncertainties on the representativeness of the results, in vitro data play a very small role in setting SMACs.

2.1.3.2. Species as a Selection Criterion

Publications dealing with the toxic effects on plants and nonmammalian animals are usually ignored because the meaning of a toxic effect in plants or nonmammalian animals is doubtful for astronauts, the subjects protected by the SMACs.

2.1.3.3. Exposure Route as a Selection Criterion

In terms of the exposure route, there is little emphasis on studies done using oral or cutaneuous routes or studies in which the chemical was administered into the experimental subject by injection. This approach is taken because, to estimate a safe exposure limit for an airborne chemical, the toxic effects by inhalation exposures are most relevant. Although there are ways to estimate the equivalent inhalation exposure concentrations for exposures by noninhalation routes, there are controversies on the accuracy of the methods. Some toxicologists even take the position that extrapolation of toxicological data from a noninhalation route of exposure to inhalation is invalid. However, if there is no toxicity information on a chemical given by inhalation and a "safe" airborne exposure limit is definitely needed, then sometimes we may calculate an inhalation exposure concentration that is somewhat equivalent to a dose of the chemical given by injection or oral administration. The calculation assumes that the chemical is absorbed to the same degree regardless whether the exposure is by an oral or inhalation route and that the absorption rate stays constant during the entire duration of an inhalation exposure. The equivalent inhalation exposure concentration is calculated using the typical minute volume of an average adult. Even though these assumptions may not be precisely correct, they tend to err on the conservative side, so the extrapolation method is considered an acceptable means of calculating the equivalent inhalation exposure concentration as a last resort.

2.1.4. Other Types of Information Gathered

In addition to reviewing the literature on the toxicity of the chemical, the literature on the absorption, metabolism (how the body modifies the chemical), excretion, pharmacokinetics (how fast the body handles the chemical), and mechanism of toxicity (how the chemical causes toxic effects in the body) is scrutinized. Such data are important in setting the SMACs, but they are typically incomplete.

Absorption and excretion information is useful. Knowing

how well the body takes in the chemical and how well it eliminates the chemical provides a better understanding of the biological fate of the chemical, which in turn helps in the interpretation of the toxicity data.

Metabolism information is important because some chemicals are toxic after the body converts the chemicals into metabolites which are more toxicologically active than the parent chemicals. For these chemicals, the SMACs should be set according to the potential level of the active metabolite in the body and not according to that of the parent chemical.

Knowing the pharmacokinetics of the chemical could provide us a way to estimate the level of the chemical at the target site in the body at different times during or after an inhalation exposure. Because the toxicity depends on the chemical level at the target site, the estimates of that level with time allows the proper adjustments of toxicity data from a certain exposure duration in a study to the duration for which the SMAC is designed.

2.2. Evaluation of the Toxicological Literature

Evaluation of the information on toxicity, metabolism, pharmacokinetics, excretion, and mechanism of toxicity provides a thorough understanding of the toxic nature of the chemical in the biological system. They are of value in making predictions on the nontoxic exposure concentrations for different exposure durations.

2.2.1. Species

Because a chemical's toxicity, metabolism, pharmacokinetics, excretion, and mechanism of toxicity could vary depending on the species of the exposed subject, it is important that SMACs should be set relying on data gathered from human subjects whenever possible. Therefore, in evaluating the toxicological literature, emphasis is placed on human studies. Due to ethical reasons, many types of toxicological studies are not conducted on human subjects. As a result, we sometimes have to rely on toxicological data gathered in studies using laboratory animals. Among the data gathered from laboratory animals, more weight is given to data in nonhuman primates because of the physiological and anatomical similarities of monkeys, baboons, etc., to man. For data gathered in nonprimate animal species, a slight preference is given to larger animal species, e.g., the dog, than rodents. The reason is that the human physiology and anatomy are more analogous to that in larger species.

2.2.2. Exposure Duration

Because the pattern of toxicity of a chemical could vary with the exposure duration, the toxicity data are categorized into acute, subchronic, and chronic. Acute toxic changes are adverse effects produced by the chemical after a single, short-term exposure. Adverse effects on the body caused by the chemical after an exposure lasting almost a normal life span

614

I

are termed chronic toxic changes. Subchronic toxicity is the toxic effect produced by a chemical exposure that lasts between a short-term exposure and a near life-time exposure. Actually, the transition from acute toxicity to subchronic toxicity sometimes is imprecise because the definition of subchronic toxicity is not very specific. Therefore, some toxicologists have added a quantitative criterion into the definition. They define subchronic toxicity as the toxic effect seen after the subject has been exposed to the chemical for less than a life time, but more than 10% of the life time. Because 10% of a life span of 70 year is too long a time relative to the exposure durations of the SMACs, here we have chosen the less specific definition of subchronic toxicity. We have opted for a definition that any exposure lasting for about a week to almost a life time is subchronic.

Because the toxicity of a chemical could differ according to the exposure duration, the categorization of the toxicity information found in the literature into acute, subchronic, and chronic types is important for the setting of SMACs of different duration. Ideally, a SMAC should be set using toxicological data generated in a study in which the exposure duration was similar to the exposure duration that the SMAC is intended to be used. So the 1-hr and 24-hr SMACs should be set using acute toxicological data. By the same token, subchronic toxicological data should be used to establish the 7-day, 30-day, and 180-day SMACs.

Chronic toxicological data usually yield information on the carcinogenicity of the chemicals. Tumors are thought to be produced via a different process than the production of noncarcinogenic endpoints, e.g., liver injury. The setting of SMACs based on the carcinogenesis of a chemical, consequently, involves a different approach than the setting of SMACs based on noncarcinogenic endpoint. Such an approach will be described separately below.

2.2.3. Data Quality

To make the SMACs meaningful, it is obvious that the SMACs should be established based on good quality data. In evaluating the literature, attention is always placed on the data quality. The factors considered in assessing the data quality are the number of test subjects, the homogeneity of test subjects, the way the exposure concentration was determined in the study, the purity of the chemical used, the method of detecting the toxic changes, and the way study results were analyzed.

2.2.3.1. Number of Test Subjects

The study has to use a sufficient number of test subjects in order to make the response detected in the study representative of what would be seen in a population. The sample size is important because, unlike physical or chemical systems, biological systems could have quite a large degree of variability due to differences in the genetic makeup. If the data were gathered in a study with only a few human subjects,

the average response in these subjects most probably would not be predictive of how other human beings, such as astronauts, would react to the chemical. As a result, SMACs set on the results of such a study are of doubtful value.

2.2.3.2. Homogeneity of Test Subjects

While the homogeneity of the test subjects used would not, by itself, eliminate a study from being relied upon in setting SMACs, it is still an important criterion. Because a test subject's response to a chemical is influenced by many factors related to the subject, it is important that the study subjects are very similar anatomically and physiologically so that any changes seen in the study could be ascribed to the chemical exposure alone. Therefore, we prefer studies in which the composition of the exposed group is very similar to that of the control (nonexposed) group. Confounding factors, such as existing disease state and smoking habit, that may influence the response to the chemical, should be eliminated to the extent possible. Given the impossibility of conducting a study with a perfectly homogeneous human population, these confounding factors should be randomized evenly between the exposed and the nonexposed group. In laboratory animal studies, the control of confounding factors is easier. So among animal studies, an emphasis should be placed on studies with healthy animals of similar age and body weight.

2.2.3.3. Measurement of Exposure Concentrations

Since the SMACs are set based on the toxic effects detected at some specific exposure concentrations in studies in the literature, it is important that the reported exposure concentrations were accurate. Consequently, during the literature evaluation, attention is paid to the concentration measurement method. Questions usually asked include whether the exposure concentration was measured analytically or estimated from the amount of chemical introduced into the exposure system. If the concentration was measured analytically, we would determine if the analytical method is specific for the chemical and if it is sensitive enough in the range of the study concentrations.

2.2.3.4. Chemical Purity

An issue related to the method of determining the exposure concentration is the purity of the chemical used in the study. To eliminate exposures to chemicals other than the study chemical as a confounding factor, preference is given to studies in which the subjects were exposed to only one chemical. That means in experimental studies, the study chemical used should be of high purity (over 99% preferably and at least 95%). Attention should be paid to the identities and relative quantities of the impurities. In epidemiological or occupational studies, it is difficult to control the exposure, so there is no exact purity requirement. For this reason, experimental human studies are preferred over epidemiological or occupational studies.

2.2.3.5. Detection Method of the Toxic Effect

Another important evaluation criterion is the method used to detect the toxic effect. Other than studies investigating chemical-induced headache and mucosal irritation, objective methods are given more weight than subjective methods to avoid biases of the investigators. If the study involves human subjects, it is prefered that the subjects did not know whether they were being exposed to the chemical or just air. If the method of detecting the toxic effect involves the investigators' judgment, a double blinded study is prefered. Other factors considered include whether the method is specific for the toxic effect to be detected, the method's sensitivity and accuracy, and the reproducibility of the method. It is obvious that the specifity and accuracy of the endpoint detection method are important because the measured change should reflect the real toxic effect. By the same token, any lack of measurable changes should represent a real absence of toxic effect, so the detection method must be sensitive. Finally, the method's reproducibility is important in ascertaining that the measured change was caused by the chemical and not due to any aberrant behavior of the detection method.

2.2.3.6. Data Evaluation in the Study

To ensure that the investigators drew the correct conclusion in the study, it is important that the data were analyzed properly. To deal with biological data that could vary tremendously, the use of statistics is imperative. We routinely check whether the toxic changes reported were based on suitable statistical analyses.

3. SMAC Setting Procedure

The general procedure is to set a maximum allowable concentration (MAC) for each toxic endpoint meaningful for the exposure duration of interest and select the lowest MAC, among those of different endpoints, as the SMAC for that duration. The way to set a MAC for carcinogenesis is different from that for noncarcinogenic endpoints. The ways for setting the two types of MACs are described separately.

3.1. MAC based on noncarcinogenic endpoint

The SMAC for an exposure duration is defined as the maximum exposure concentration that causes no or practically no toxicity for that exposure duration. The MAC for a particular toxic endpoint is, then, the maximum exposure concentration that would not cause that toxic effect. It is believed that the higher the exposure concentration the greater the toxicity. So, from the study results, we are interested in the maximum exposure concentration that failed to produce a specific toxic effect to a practical extent and such a concentration is called a no-observed-adverse-effect-level or NOAEL. To obtain the NOAEL for a noncarcinogenic endpoint, the acute, subchronic, and chronic toxicity data are

separated. In each of these three categories, the data are further separated according to the toxic endpoint. The highest exposure concentration known not to cause a particular noncarcinogenic toxic effect is identified as the NOAEL for that effect.

3.1.1. Safety Factors

Due to uncertainties inherent in some studies, to err on the conservative side, the NOAEL for a particular noncarcinogenic endpoint is sometimes divided by safety factors to arrive at the MAC. If the study used animals as the test subjects, an interspecies extrapolation factor may be used. If the study was of an exposure duration different from the exposure duration of the SMAC, a time-adjustment factor is applied. The time-adjustment factor is based on the Haber's rule that the product of the exposure concentration and exposure time is approximately constant. It should be pointed out that the Haber's rule is not universally applicable.

Another situation in which a safety factor is applied is when the chemical's toxic effect is similar to a biological effect induced by microgravity. For instance, microgravity has been known to reduce the red blood cell mass and to cause minor irregular heart rhythm in astronauts during a mission[7,8]. If a chemical also produces a change in the red blood cell or heart rhythm, a microgravity-related safety factor is usually applied on the NOAEL to derive the MAC.

3.2. MAC based on carcinogenesis

Most carcinogens have been shown to act by damaging the genetic material of the body and these are called genotoxic carcinogens. Because there is a better understanding of genotoxic carcinogens, the setting of MAC based on carcinogenesis is described using genotoxic carcinogens as examples. The current thought is that there is no such thing as a NOAEL for the tumor response to genotoxic carcinogens. That means even a single molecule of genotoxic carcinogen in the body could theoretically lead to tumor development. The JSC Toxicology Group has adopted 1/1000 as an acceptable tumor risk. So the maximum concentration of a carcinogen that could produce one case of tumor, in excess of the background, in 1000 exposed individuals is the MAC for that carcinogen.

The MAC based on carcinogenesis is usually generated from chronic test data in laboratory animals. Using these data, the U.S. Environmental Protection Agency or the NRC-COT calculated, with the linearized multistage model, the upper 95% confidence limit of the life-time exposure concentration that yields an excess tumor risk of 1/1000. We obtain the MAC based on carcinogenesis by taking that upper confidence limit and adjust, with the method of Crump and Howe[9], for the difference in exposure duration between the continuous, life-time exposure of the EPA or NRC-COT value and the exposure duration for which the SMAC is designed.

4. Setting the SMACs for Carbon Monoxide

The setting of the carbon monoxide's SMACs will be used to illustrate the procedure of establishing SMACs described above.

4.1. General Biological Information on CO

Carbon monoxide is absorbed rapidly in the lung at a rate of about 25.8 ml /min x mm Hg[10]. Carbon monoxide's elimination half-life is 4-5 hours in resting subjects when breathing ambient air[4].

As mentioned earlier, an understanding of the mechanism of toxicity is very useful in setting SMACs. Information on the mechanism of toxicity of carbon monoxide indicates that, unlike most chemicals, the exposure concentration of carbon monoxide together with the exposure duration is not the best way to characterize carbon monoxide's toxic potential. Carbon monoxide acts by reversibly binding to the heme group in hemoglobin, forming carboxyhemoglobin (COHb)[12]. As a result, less oxygen will be carried by blood to tissues during carbon monoxide inhalation. That means organs, such as the heart and the brain, having a critical need for oxygen would be the major target organs for carbon monoxide poisoning. Indeed, data indicate that carbon monoxide's toxicity is manifested primarily on its effects on the heart and the brain[4]. However, it should be pointed out that not all COHb levels are harmful to the health. For instance, the normal metabolism of heme generates carbon monoxide inside the body, leading to a COHb level of 0.4-0.7%, which does not result in adverse effects[4].

Due to the fact that carbon monoxide's toxicity is correlated with the COHb level in the blood, the SMACs were set according to safe COHb levels and the toxicity data were reported in terms of the COHb levels. The COHb level will rise with time as an individual inhales carbon monoxide and it would reach a plateau within 24 hours[13]. The resulting COHb level of a carbon monoxide can be calculated using the Coburn-Forster-Kane (CFK) equation[13].

4.2. Toxicity Information on Carbon Monoxide

The most important information for the establishment of SMACs is the information on the toxic effects of the chemical. To facilitate the setting of short-term and long-term SMACs, toxicity information on carbon monoxide was separated into three categories: acute, subchronic, and chronic.

4.2.1. Acute Toxicity

This subsection summarizes the toxicity of carbon monoxide in humans after an exposure ranging from several minutes to 8 hours. A COHb level of 70% is lethal, while 50-60% leads to coma and convulsions[4]. Headache, nausea, and vomiting are detected are detected at 15-40% COHb[14,15]. Increased

frequencies of ventricular premature depolarization were seen in coronary patients at 6% COHb[16]. A COHb level of 5% is known to increase reaction time and impair hand-eye coordination[11,17]. Finally, carbon monoxide reduces the maximal exercise duration at 3.4% COHb[18].

4.2.2. Subchronic Toxicity

This subsection summarizes the toxicity produced by carbon monoxide exposures lasting longer than one day. A continuous exposure of humans to carbon monoxide for 7 or 8 days resulting in 2.4% COHb has been shown to produce heart rhythm changes[19]. New York employees working in a tunnel with, on the average, 38 ppm carbon monoxide, as well as other automobile exhaust pollutants, showed a higher mortality from arteriosclerotic heart disease[20]. However, whether subchronic carbon monoxide exposures cause arteriosclerotic heart diease is still not certain.

There are a lot of data on carbon monoxide's subchronic toxicity gathered in laboratory animals, but only the more important data will be described here. As discussed above, carbon monoxide intoxication primarily produces toxicity on the brain and the heart. However, animal studies demonstrated that, in subchronic exposures, carbon monoxide could also cause increases in hemoglobin concentration in the blood. For instance, a continuous, 6-month exposure of dogs to 50 ppm carbon monoxide, which is equivalent to 7.3% COHb, increased the hemoglobin concentration by 12%, but it did not cause any changes in the EKG[21]. Similarly, a continuous, 90-day exposure of rats to 96 ppm carbon monoxide (equivalent to 7.5% COHb) increased both the hemoglobin concentration and the hematocrit[22].

The increases in hemoglobin and red blood cells probably represent an adaptation of the body to the subchronic deficiency of oxygen supply to tissues. It should be noted, however, that there are species differences in this adaptive response. No increases were seen in monkeys exposed to 66 ppm of carbon monoxide, reaching a COHb level of 7.4%, for 2 years[23]. It took a much higher level of carbon monoxide exposure for such an adaptive response to occur in the monkey. A continuous exposure of monkeys to 200 ppm carbon monoxide (equivalent to 16 to 20% COHb) increased the hematocrit and hemoglobin level[22]. Because humans are biologically closer to monkeys than to the dog or rat and because astronauts are not expected to encounter high levels of carbon monoxide for a long time, this adaptive response of the body toward subchronic carbon monoxide most likely would not occur in the astronauts. As a result, the adaptive response was not considered in setting the carbon monoxide SMACs.

4.2.3. Chronic Toxicity

There are no data on the toxic effects of a carbon monoxide exposure lasting almost the normal life span of an adult.

4.3. Rationale Used to Set the Short-term SMACs

In an acute carbon monoxide exposure, the lowest COHb level that has been found to adversely affect the body is 3.4%, which reduced the maximal exercise duration by 5% in human volunteers[18]. Since a mere 5% reduction in the maximal exercise duration would not significantly impair the astronauts in an emergency, the exercise effect of 3.4% COHb is considered acceptable in short term. Therefore, the second lowest toxic COHb level is chosen to be the criterion for setting the 1-hour and 24-hour SMACs.

A short-term exposure to carbon monoxide, yielding a COHb level of 5%, is known to cause CNS impairment, such as an impairment of hand-eye coordination and an increase in reaction time[17]. The short-term SMACs should be set to prevent CNS impairment because CNS impairment would interfere with the astronauts' ability to deal with an emergency. The task is to estimate a COHb level that would not cause CNS impairment. Because 5% COHb appears to be the threshold for causing CNS impairment in humans[11] and because an 1-hour exposure to carbon monoxide at the 1-hour National Ambient Air Quality Standard would yield a COHb level of about 3% in exercising subjects[6], the 1-hour and 24-hour SMACs are set at a level that would produce no more than 3% COHb in the astronauts. Using the hemoglobin concentration determined inflight during the Skylab missions[24], the 1-hour and 24-hour carbon monoxide exposure concentrations that would yield a 3% COHb were calculated with the CFK equation[13]. The 1-hour and 24-hour exposure concentrations were calculated to be 55 ppm and 20 ppm, respectively. The 1-hour and 24-hour SMACs were set at 55 ppm and 20 ppm, respectively.

4.4. Rationale Used to Set the Long-term SMACs

In setting the 7-day SMAC, the toxicity data from an 8-day carbon monoxide exposure of human subjects were utilized[19]. In that study, an 8-day exposure to 15 ppm carbon monoxide resulted in 2.4% COHb and caused P wave changes in the EKG. Therefore, the 7-day SMAC should be established to prevent any EKG changes induceable by carbon monoxide. It has been calculated that an 8-hour exposure of exercising individuals to the EPA's 8-hour National Ambient Air Quality Standard would yield a COHb of 1.6%[6]. It is reasoned that if 1.6% COHb is safe for the general population, 1.6% COHb should also be safe for astronauts. The 7-day SMAC is the carbon monoxide exposure concentration that would yield 1.6% COHb in a 7-day exposure. That SMAC was calculated to be 10 ppm, applying the Skylab hemoglobin concentration in the CFK equation.

The use of the hemoglobin concentration determined during the Skylab missions in the 7-day SMAC calculation corrects for the microgravity-induced reduction in red cells. Even though cardiac arrhythmias were detected in the astronauts during the Skylab mission[25], no correction was made for the potential synergistic effects of carbon monoxide and

microgravity on the heart. The reason is that the 7-day SMAC is based on the 8-hour National Ambient Air Quality Standard, which was designed to protect the general population and the astronauts are much healthier than the infirmed or the aged in the general population.

There are no human data on carbon monoxide's toxicity for an exposure lasting longer than 8 days. The subchronic data gathered in laboratory animals[21-23] were not relied upon in setting the 30-day and 180-day SMACs because the most sensitive endpoints, i.e., CNS impairment and cardiac arrhythmias, were not reliably determined in these studies. The target COHb level of the 7-day SMAC of 1.6% was chosen to be the target COHb level for the 30-day and 180-day SMACs. A COHb level of 1.6% should provide a sufficient safety margin for the astronauts in a 30- or 180-day carbon monoxide exposure. The reason is that 1.6% is not much higher than the COHb level of 0.7-1.0% measured in nonsmokers[26] and it is also lower than the COHb level of 3% or 4-5% in smokers[27,28]. With the CFK equation, the 30-day and 180-day SMACs were calculated to be both 10 ppm[13].

5. CONCLUSION

After reviewing information on the toxicity, absorption, metabolism, and pharmacokinetics of a chemical, the toxicologist must systematically sort the data and identify the toxic endpoints that a SMAC must protect against. By applying toxicological principles to the appropriate data on the toxic endpoints, the safe exposure concentration is then estimated for a given exposure duration and such a concentration is the SMAC for that exposure duration. This paper demonstrates the approach using carbon monoxide as an example.

6. REFERENCES

1. Office of Space Transportation Systems, "Flammability, Odor, and Offgassing Requirements and Test Procedures for Materials in Environments that Support Combustion", NHB 8060.1B, NASA, Washington, DC, September 1981.

2. NRC's Committee on Toxicology, "Emergency and Continuous Exposure Guidance Levels for Selected Airborne Contaminants", Vol. 4., National Academy Press, Washington, D.C., 1985, p. 17.

3. Coburn, R.F., et al., "Considerations of the physiology and variables that determine the blood carboxyhemoglobin concentration in man", J. Clin. Invest., Vol. 44, 1964, pp. 1899-1910.

4. Stewart, R.D., "The effect of carbon monoxide on humans", Ann. Rev. Pharmacol, Vol. 15, 1975, pp. 409-423.

618

5. Lawther, P.J., "Carbon monoxide", Br. Med. Bull., Vol. 31, 1975, pp. 256-260.

6. Rylander, R. & Vesterlund, J., "Carbon monoxide criteria", Scand. J. Work Environ. Health, Vol. 7, Suppl. 1, 1981, pp. 1-39.

7. Huntoon, C.L., et al., "Hematology, immunology, endocrinology, and biochemistry", SPACE PHYSIOLOGY AND MEDICINE, Ed. Nicogossian, A.E., Lea & Febiger, Philadelphia, PA, 1989, pp. 222-239.

8. Bungo, M., "The cardiopulmonary system", SPACE PHYSIOLOGY AND MEDICINE, Ed. Nicogossian, A.E., Lea & Febiger, Philadelphia, PA, 1989, pp. 179-201.

9. Crump, K.S. & Howe, R.B., "The multistage model with a time-dependent dose pattern: Applications to carcinogenic risk assessment", Risk Anal., Vol. 4, 1984, pp. 163-176.

10. Jones, H.A., et al., "Rate of uptake of carbon monoxide at different inspired concentrations in humans", J. Appl. Physiol., Vol. 52, 1982, pp. 109-113.

11. Ramsey, J.M., "Carbon monoxide, tissue hypoxia, and sensory psychomotor response in hypoxaemic subjects", Clin. Sci., Vol. 42, 1972, pp. 619-625.

12. Laties, V.G. & Merigan, W.J., "Behavioral effects of carbon monoxide on animals and man", Ann. Rev. Pharmacol., Vol. 19, 1979, 357-392.

13. Peterson, J.E. & Stewart, R.D., "Predicting the carboxyhemoglobin levels resulting from carbon monoxide exposures", J. Appl. Physiol., Vol. 39, 1975, pp. 633-638.

14. DiMarco, A., "Carbon monoxide poisoning presenting as polycythemia", N. Engl. J. Med., Vol. 318, 1988, p. 874.

15. Stewart, R.D., et al., "Experimental human exposure to carbon monoxide", Arch. Environ. Health, Vol. 21, 1970, pp. 154-164.

16. Sheps, D.S., et al., "Production of arrhythmias by elevated carboxyhemoglobin in patients with coronary artery disease", Ann. Intern. Med., Vol. 113, 1990, pp. 343-351.

17. Putz, V.R., et al., "A comparative study of the effects of carbon monoxide and methylene chloride on human performance", J. Environ. Pathol. Toxicol., Vol. 2, 1979, pp. 97-112.

18. Horvath, S.M., et al., "Maximal aerobic capacity at different levels of carboxyhemoglobin", J. Appl. Physiol., Vol. 38, 1975, pp. 300-303.

19. Davies, D.M. & Smith, D.J., "Electrocardiographic changes in healthy men during continuous low-level carbon monoxide exposure", Environ. Res., Vol. 21, 1980, pp. 197-206.

20. Stern, F.B., et al., "Heart disease mortality among bridge and tunnel officers exposed to carbon monoxide", Am. J. Epidemiol., Vol. 128, pp. 1276-88.

21. Musselman, N.P., et al., "Continuous exposure of laboratory animals to low concentration of carbon monoxide", Aerosp. Med., Vol. 30. 1959, pp. 524-529.

22. Jones, R.A., et al., "Effects on experimental animals of long-term inhalation exposure to carbon monoxide", Toxicol. Appl. Pharmacol., Vol. 19, 1971, pp. 46-53.

23. Eckardt, R.E., et al., "The biologic effect from long-term exposure of primates to carbon monoxide", Arch. Environ. Health, Vol. 25, 1972, pp. , 381-387.

24. Kimzey, S.L., "Hematology and immunology studies", BIOMEDICAL RESULTS FROM SKYLAB, Ed. R.S. Johnson & L.F. Dietlein, 1977, NASA, Washington, D.C., p. 249.

25. Smith, R.F., et al. (1977). "Vectorcardiographic changes during extended space flight (M093). Observations at rest and during exercise", BIOMEDICAL RESULTS FROM SKYLAB, Ed. R.S. Johnson and L.F. Dietlein, 1977, NASA, Washington, D.C., pp. 339-350.

26. Radford, E., et al., "Blood carbon monoxide levels in persons 3-74 years of age in the United States, 1976-80", Advance Data Report No. 76, 1981, National Center for Health Statistics, Hyattsville, MD.

27. O'Hanlon, J.F., "Preliminary studies of the effects of carbon monoxide on vigilance in man", BEHAVIORAL TOXICOLOGY, Ed. B. Weiss & G. Laties, 1975, Plenum Press, N.Y., pp. 61-75.

28. Horvath, S.M., et al., "Maximal aerobic capacity at different levels of carboxyhemoglobin", J. Appl. Physiol., 1975, pp. 300-303.

7. ACKNOWLEDGEMENT

HUMAN EXPOSURE LIMITS TO HYPERGOLIC FUELS

H. D. Garcia*, J. T. James, and T. F. Limero*
NASA Biomedical Operations and Research Branch, SD4, Johnson Space Center, Houston, TX 77058
and *Krug Life Sciences, 1290 Hercules Drive, Suite 120, Houston, TX 77058

ABSTRACT
Over the past four decades, many studies have been conducted on the toxicities of the rocket propellants hydrazine (HZ) and monomethylhydrazine (MH). Numerous technical challenges have made it difficult to unambiguously interpret the results of these studies, and there is considerable divergence between results obtained by different investigators on the inhalation concentrations and exposure durations required to produce a given toxic effect. To determine the safe maximum acceptable concentrations (MACs) for each toxic effect inducible by exposure to hypergolic fuels in spacecraft atmospheres, NASA undertook a critical review of published and unpublished investigations on the toxicities of these compounds. The quality of the data from each study was assessed based on current state-of-the-art practices for similar studies. While many questions remain unanswered, MACS were determined using the best available data for a variety of toxic endpoints for potential continuous exposure durations ranging from 1 hour to 180 days. Spacecraft MACs (SMACs) were set for each compound based on the most sensitive toxic endpoint at each exposure duration.

INTRODUCTION
Payload and utility chemicals to be used or generated on manned space flights are evaluated by toxicologists at NASA for their potential health effects on crew members. For each mission, toxicologic information summaries and brief hazard assessments for each chemical flown are published as a "Blue Book". In addition, SMACs are established for individual priority airborne chemicals by NASA and contractor toxicologists after in-depth toxicological evaluations of the chemical. The SMAC documents include summaries of the chemical, physical, pharmacokinetic, and toxicological properties of each chemical and document the rationale used to establish the SMAC values. The SMAC values are used as guidelines to determine the levels of containment needed for each chemical flown and the health impact of a chemical release into the cabin atmosphere.

NASA has a policy to comply with maximum exposure limits for the workplace (ground based operations) established by other nationally recognized groups such as the American Congress of Governmental Industrial Hygienists (ACGIH) Threshold Limit Value (TLV®) Committee. The ACGIH, for example, promulgates recommended exposure limits (TLVs) for airborne chemicals in the workplace atmosphere, based on the typical occupational exposure pattern, i.e. 8 h/day, 5 d/wk, for a working lifetime. This exposure pattern is quite different from that experienced by astronauts, i.e. continuous exposures for periods of up to

one week, typically, and eventually up to 6 months or longer. The TLVs, thus, are applicable for many ground-based NASA operations. The recent announcement by the ACGIH's TLV committee proposing to lower the TLV for HZ from 0.1 ppm to 0.01 ppm generated interest within NASA and the U.S. Air Force due to the difficulty of rapidly measuring 0.01 ppm. The stated basis for lowering the TLV ten fold was a study[1] which showed an increase in nasal tumors at 50 ppb. However, the authors of the cited study stated in their report that this result was not statistically significant. At this same time, JSC toxicologists were preparing SMAC documents on both HZ and MH. Based on their review of the literature on HZ, some felt that the scientific literature did not support lowering of the TLV value. While safety should never be compromised, we must also take into account the fact that limited resources would be wasted if MACs for hypergolic fuels were set at unnecessarily low values. A preliminary review of the impact of lowering the current standards for safe levels of these fuels indicates that important NASA and Air Force operations could be restricted due to the need to:

1) revise current operating procedures for determining and certifying that the work environment contains less than 10 ppb of either hydrazine,

2) develop instruments with enough sensitivity to measure very low concentrations of fuels in real time, and

3) improve engineering controls to restrict exposure of workers.

The following is a description of the process that the JSC toxicologists used to evaluate the toxicological data available in the literature for HZ and MH and to establish SMAC values for these compounds.

PROCEDURE FOR SETTING SMAC VALUES

The information gathering phase of the SMAC setting process involves a comprehensive review of the published literature on the toxicity of the compound. This includes a search of the computerized databases (e.g. Toxline, Toxlit, Medline, etc.) for all articles on the toxicity and/or chemistry of the compound. Abstracts of all relevant articles are obtained and reviewed. The key articles are identified and complete copies of the original articles are obtained whenever possible. A toxicity table is compiled containing a summary of the toxicity findings for inhalation exposures, including species exposed, exposure durations and chemical

concentrations, and toxic effects observed. If insufficient data is available from inhalation exposures, toxicity data from other routes of exposure must be used.

The evaluation phase involves a determination of the biological endpoints of most concern and a critical review of the quality of the available data. The criteria used to rank the toxic endpoints include the severity of immediate effects on crew performance, which could jeopardize mission-related activities, and the severity of long term health effects on crew members. The criteria used to judge the quality of the studies include such factors as:

Was the chemical pure and was its concentration determined analytically at frequent intervals during exposure?
Was the exposure chamber operated according to accepted guidelines?
Were there confounding exposures to other agents?
Were a sufficient number of subjects exposed?
Were there sufficient numbers of appropriate controls which were properly sham exposed?
Were all the subjects & controls healthy (other than treatment-related effects)?
Were relevant endpoints examined after appropriate exposure periods?

Based on these criteria, the highest quality, relevant studies are identified and used as starting points to derive SMAC values.

The extrapolation phase of the SMAC setting process involves extrapolation of data from the key studies to relevant human exposure conditions to obtain MACs for each endpoint of concern at exposure durations of 1 hr, 24 hr, 7 days, 30 days and 180 days. Thus, each exposure duration may have multiple MACs, each based on one biological endpoint such as carcinogenesis, irritation, central nervous system effects, lethality, cardiovascular effects, etc. The use of adjustment factors is often required to account for differences in species, routes of administration, exposure duration, and effects of microgravity. Extrapolation to untested exposure durations is often based on the unproven, but widely used assumption that the toxicity of a compound is a linear function of the product of the exposure duration, t, and the concentration of the compound, c. If the product of c times t for various exposure scenarios is equivalent, then equivalent toxic effects can be expected. This method of extrapolation is not universally applicable, however, and must be used with caution.

The value of the MAC for the most sensitive endpoint is selected for each exposure duration to establish SMAC values. The rationale used to set these "INTERIM" SMAC values is documented along with the information used from the key studies.

The review phase of the SMAC setting process involves sending the INTERIM SMAC document to the National Research Council's Committee on Toxicology for comments. After return of the document by the NRC's COT, it is revised as needed and an OFFICIAL SMAC established.

The following descriptions illustrate how the process described above was applied in establishing SMAC values for HZ and for MH.

HYDRAZINE

A computerized search of several toxicology-related databases identified numerous articles dealing with HZ and related compounds. Of these, abstracts were reviewed for 65 articles whose titles indicated they would be of relevance to setting SMAC values. Full copies were obtained for 29 of these articles. Two of these articles proved to be key articles on which the SMAC values were based.

HZ vapor is extremely irritating to the eyes, nose, and throat. Quantitative worker exposure information, however, is not available and/or cannot be estimated from the existing published data[2]. The median concentration of HZ detectable by odor is 3-4 ppm[3]. Inhalation can cause dizziness, anorexia[4] and nausea. HZ can be absorbed through the skin[5] or orally and can induce contact dermatitis[6,7,8], neurological impairment[9, 10], and at a dose of 10 mg/kg injected intraperitoneally for three days in rodents, HZ induces kidney, lung, and liver damage[11]. Accidental human exposures to high, but unspecified doses induced temporary blindness, and tremors[12] and, in one case involving 6 months of occupational contact exposure, conjunctivitis, tremor, cough, fever, vomiting, diarrhea, and death[13]. A two hour exposure to 1.0 to 2.0 mg/l (7600 to 15200 ppm) of HZ vapors has been reported to induce convulsions, respiratory arrest and death in mice and rats[10]. The toxicity of multiple lower doses was cumulative, but surviving animals recovered and lived normal lifespans if exposure was discontinued. The LC_{50} for a four hour exposure to HZ was 750 mg/m^3 for rats and 330 mg/m^3 for mice[3]. Exposure to 1.0 ppm of HZ for 90 days was highly lethal to rats and mice and moderately lethal (20%) to monkeys[14]. HZ was fetotoxic in rats and mice at 5 mg/kg administered intraperitoneally and was teratogenic in mice at 12 mg/kg. HZ was mutagenic in several test systems[15, 16, 17] and induced sister chromatid exchanges in vitro [18].

There are conflicting reports on HZ's carcinogenicity. The National Institute of Environmental Health Sciences finds that there is sufficient evidence for the carcinogenicity of HZ in experimental animals, but inadequate evidence for its carcinogenicity in humans[19]. Inhaled HZ has been reported to induce alveolargenic carcinomas in three of eight mice exposed at 0.2 ppm[20], nasal and bronchial tumors in rats at 1 ppm[21] (see Table 1 below), and nasal tumors in 16 of 160 hamsters exposed at 5 ppm[21]. Unequivocally toxic doses (up to 50 mg/l), however, administered in drinking water for the lifetime of rats were only weakly carcinogenic[22]. In male rats (the most sensitive species and sex) exposed by inhalation, tumors (which were predominantly benign) appeared only late in life in animals showing many other chronic toxic effects including a greatly increased inflammatory response of the upper airways[21]. An additional study would be extremely helpful. Only scanty epidemiological data is available on HZ-related cancers in humans[2]. No excess risk of cancer has been found in workers occupationally exposed to HZ vapors, but the number of workers having substantial exposure has been too few to detect anything other than gross hazards[23, 12].

Of the many toxic effects of HZ exposure, the ones of most concern, for which quantitative data are available include hepatotoxicity, carcinogenicity, and lethality. These effects were evaluated and MACs were set for each endpoint at the various exposure durations.

Hepatotoxicity:

Liver effects induced by airborne HZ include focal liver cell

621

hyperplasia in female rats exposed at 1 ppm HZ for 2 yrs and "fatty changes" in the liver in several species after 90 days continuous exposure to 0.8 ppm HZ[25, 14]. MAC values for liver toxicity were calculated for 180 d, 30 d, and 7 d. The use of the "concentration times time" rule to extrapolate to exposure times of 24 hrs or less (a two order of magnitude extrapolation) was deemed inappropriate for a compound such as HZ which has such a steep dose-response curve.

Carcinogenicity:

HZ has been found to be carcinogenic in animal model systems[21, 25, 17]. The oncogenic changes were mostly benign and observable only at the microscopic level, producing little or no impairment of respiratory function and no effect on life expectancy. The non-oncogenic toxicities of HZ exposure in animals were more severe in producing debilitation and lethal effects. There are, moreover, no known reports of HZ-induced human tumors. Most human exposures to HZ have been accidental or job-related and dose-response data are not available. Although the data for a carcinogenic effect of HZ are not compelling, prudence requires that carcinogenicity be considered in determining a MAC for HZ. MACs were calculated using the linearized multistage model method described by the NRC[27] as illustrated below.

Based on the data of MacEwen et al., the NRC Committee on Toxicology (COT) calculated that the lower 95% confidence limit on the inhalation dose that would produce a 1% lifetime tumor incidence in rats is 0.055 ppm for a 6 hr/day, 5 day/wk, 52 wk/yr, 1 yr exposure. This extrapolates to 0.005 ppm for a continuous 2 yr exposure:

0.055 ppm x (6 hr / 24 hr) x (5 hr / 7 hr) x (1 yr / 2 yr) = 0.005 ppm

Extrapolating the 1% tumor incidence from a continuous 2 year exposure at 0.005 ppm to a 24 hr exposure at the same 0.005 ppm concentration, the NRC COT estimated[28] the tumor risk for rats should be no more than 10^{-2}.

The NRC has stated[29] that the linearized multistage model is sufficiently conservative so that an additional species extrapolation factor is unnecessary. Therefore, the following equation, based on Crump and Howe's linearized multistage model[27], was used to calculate the exposure concentrations, D, which would yield a tumor risk of 10^{-3} for exposure durations of 1 h, 24 h, 7 d, 30 d, and 180 d.:

$$D = (d) \ (25,600)^k \ (10^{-3}/\text{risk}) \div \{25,600 - ((365)(\text{age}))^k \} - 25,600 - ((365)(\text{age}) - \text{exp})^k$$

where
   d is the the concentration during a lifetime exposure
      (0.005 ppm in this case)
   $25,600$ is the # of days in a 70 y human lifetime
   k is the number of stages in the model
      (3 in this case)
   10^{-3} is the acceptable risk level
   age is the minimum age of an astronaut, in years
      (30 in this case)
   exp is the exposure duration, in days
      (0.042, 1, 7, 30, or 180)
   $risk$ is the risk of tumor for lifetime exposure to d
      (10^{-2} in this case)

This equation yields the following values, rounded to one significant figure:

300	ppm for	1 hr
10	ppm for	24 hr
2	ppm for	7 d
0.4	ppm for	30 d
0.07	ppm for	180 d

Lethality:

Analysis of the lethality data, except for very short exposures, is difficult and frustrating. To set the 1-h MAC, we rely on the LC_{50} data from an old report[8] that, nonetheless, appears to be well done. The LC_{50} for a 4-h exposure was 570 ppm in rats and 250 ppm in mice. The results for the more sensitive species (mouse) were adjusted by the following factors: 250 ppm was divided by 10 for inter-species extrapolation, again divided by 10 to extrapolate from the LC_{50} to a NOAEL, and multiplied by 2 to extrapolate from a 4-h exposure to a 1-h exposure to yield a 1-h MAC for lethality of 5 ppm. A factor of 10 was judged to be adequate to go from the LC_{50} to a NOAEL due to the steepness of the dose response curve[30]. The use of a two-fold rather than a four-fold safety factor to extrapolate from a 4 h exposure to a 1 h exposure is supported by experimental results for a similar compound, MH[30]. The 24-h MAC is calculated similarly, but dividing by 6 to extrapolate from a 4-h to a 24-h exposure to yield a 24-h MAC value of 0.4 ppm. The use of the 'concentration times time" rule for longer exposures is not considered appropriate for the available acute data. To set lethality MACs for the longer periods, we must turn to repeated intermittent exposures available on multiple species.

The scatter in the lethality data for repeated exposures to HZ is large and suggests serious shortcomings in some of the study designs. An early report[31] showed that at a nominal concentration of 20,000 ppm, the recovery of HZ decreases from 26% to 4% simply because of the presence of rat bodies, whereas, if the rats are alive, the recovery is decreased to 2%. This clearly indicates that a large fraction of the airborne HZ adheres to the rat fur, probably about 10 times the amount retained in the respiratory system. Since rats preen, much of their exposure to HZ may have been by oral ingestion, rather than inhalation. The point is that the rodent data are so scattered that the true susceptibility is simply unknown.

In the study above[31], food was removed during exposures, so that it was not contaminated with HZ. In some later studies[32], food was not removed during exposure and the rodents were found to be much more susceptible. For example, an 8% mortality was seen in mice exposed to 1.0 ppm HZ for 6 h/d for 2 wk (60 ppm•h)[32], whereas no mortality was seen in mice exposed to 40 ppm HZ 6 h/d for 4 d (960 ppm•h)[31].

Based on these studies, MACs were calculated for lethality, carcinogenesis, and hepatotoxicity as described in the procedure section.

622

TABLE 1: MACs and Endpoints for HZ

Endpoint	Species Tested	Safety Factors Species	Safety Factors NOAEL	MAC (ppm) 1 hr	24 hr	7 d	30 d	180 d
Lethality [8] $LC_{50}=250$ ppm, 4 h	Mouse	10	10	5	0.4	*	*	*
Lethality [24] 0.8 ppm, 30 d contin.	Monkey	10	10	*	*	0.08	0.02	0.003
Carcinogenesis[27] 0.05 ppm, 1 yr, contin	Rat	1	1	300[†]	10[†]	2[†]	0.4[†]	0.07[†]
Hepatotoxicity [24] 0.78 ppm, 90d, contin.	Monkey Rat, Mouse	10	10	*	*	0.1	0.02	0.004

* N/A for use of "concentration times time" rule
† Calculated based on NRC COT's equation[27] derived from Crump & Howe's multistage carcinogenicity model, using a lifetime cancer risk of 10^{-3}.

For each exposure duration, the SMAC was set equal to the lowest MAC value among the three endpoints shown above. The resulting values, in units of both ppm and mg/m³, are listed along with the most sensitive target toxicity at that exposure duration.

TABLE 2: SMACS for HZ

	ppm	mg/m³	Target Toxicity
1-h SMAC	5	13	Lethality
24-h SMAC	0.4	0.5	Lethality
7-d SMAC	0.08	0.10	Lethality, Hepatotoxicity
30-d SMAC	0.02	0.03	Lethality, Hepatotoxicity
180-d SMAC	0.003	0.004	Lethality, Hepatotoxicity

* Temporary 7-d SMAC was set at 0.04 ppm

These SMAC values, except for 180 days continuous exposure, are all above the 0.01 ppm TLV proposed by ACGIH. As will be shown below, the situation for MH is quite different and highlights the need for better, more definitive toxicological studies on this class of compounds.

METHYLHYDRAZINE

A computerized search of several toxicology-related databases identified a large number of articles dealing with MH and related compounds. Of these, abstracts were reviewed for 173 articles whose titles indicated they would be of relevance to setting SMAC values. Full copies were obtained for 79 of these articles. In contrast to the situation with HZ, the data for MH toxicity available from a study[33] performed at NASA's White Sands Testing Facility proved to be a key article which required the setting of the SMAC at a very low value.

MH can induce a variety of toxic effects. The overt signs of acute MH toxicity in mice, rats, dogs, and monkeys include irritation of nose and eyes, blood discrasias (hemolytic anemia and Heinz bodies in humans, monkeys and dogs) salivation, emesis, diarrhea, hyperactivity, tremors and severe tonic-clonic convulsions, which precede death. In addition, chronic exposure to low concentrations of MH has been shown to induce blood and liver effects in dogs and cancer in mice and hamsters.

Lethality:

A steep dose-mortality response curve was seen for all species, regardless of the length of exposure. Of the species tested, squirrel monkeys were the most sensitive to the lethal effects of MH. Monkey LC_{50} experiments were performed, examining three exposure times: 15, 30 and 60 min. and using 25 monkeys, total. The lower confidence limit was used for calculating the MAC based on the LC_{50}. This yields a value which is lower than that obtained based on the NOAEL, since no confidence limits were calculated for the NOAEL. This anomaly is due to the steepness of the dose-mortality response curve. A safety factor of 100 was used in calculating the MACs for lethality due to the severity of the endpoint.

Nasal Injury:

The most sensitive endpoint for toxicity at concentrations greater than or equal to the odor threshold was nasal injury. Since the data for injury were obtained from human subjects[32], no species conversion was required. Nevertheless, since 75% of the subjects complained of irritating odor and 28% developed significant nasal pathology under the test conditions, the tested concentration of 0.2 ppm must be lowered to a level which would be anticipated to produce no adverse effects. A safety factor of 10 was used to estimate the NOAEL. An additional safety factor of 10 was warranted to account for the fact that a single sniff caused the observed effects, whereas the MACS must be set to protect during continuous, much longer term potential exposures. The resulting level of 0.002 ppm is less than or equal to even the 180 day MACs for all other endpoints. Since the endpoint is injury of the nasal mucosa, and since no epidemiological data were available to indicate that long term exposure to sub-irritating (short-term) levels of MH would lead to cumulative effects, the MAC for nasal injury was set at 0.002 ppm MH for all exposure durations. It must be noted that the results of this NASA study do not correlate with the results reported by others for exposure of humans to MH (see Heinz bodies discussion below). Nevertheless, a careful review of the data and methodology showed that this study appeared to be well done and did not reveal any basis on which this study could be discounted or ignored.

623

Heinz Bodies:

The low level of Heinz bodies seen in human volunteers exposed to 90 ppm MH for 10 min were not enough to produce any noticeable symptoms of toxicity, and this toxicity was completely reversible(35). MACs were not calculated for exposure times longer than one hour, for the reasons stated in the footnote to the table below.

Carcinogenesis:

The NRC COT[28] used the data of Kinkead et al.[34] as input to the multistage model of Crump and Howe to obtain a 95% lower confidence limit of 0.116 ppm for a lung tumor risk in mice of 0.01, based on a work-week exposure schedule. Using the "concentrations times time" rule to convert to a continuous lifetime exposure yielded the value of 0.01 ppm corresponding to a lifetime tumor risk of 0.01.

Blood Effects:

In continuous, 90 day low dose inhalation exposures of rats, dogs and monkeys at 0.04 pm and 0.1 ppm, dogs showed significant decreases in hematocrit hemoglobin levels and blood cell count, but only at the higher dose[17]. MAC values were calculated by applying Haber's rule and a safety factor of 10 to the (NOAEL) 0.04 ppm level. Microgravity should not affect the crewmembers' susceptibility to the hemolytic effects of MH because microgravity reduces the total red cell mass, but the hematocrit remains at near normal levels because the plasma volume is also reduced.

In order to determine which of these toxic end-points should be used as the basis for setting the SMAC, MAC values were determined for each endpoint at each of five exposure durations. The results, tabulated below, were used to select the value which provided the best protection.

TABLE 3: MACs and Endpoints for MH

Endpoint		Specie Tested	Safety Factor	MAC (ppm)				
				1 hr	24 hr	7 d	30 d	180 d
Lethality[30] LC$_{50}$, 1 hr = 82 ±16 ppm (2 died/4 exposed)	Severe	Monkey	100	0.65	0.03	*	*	*
Lethality[30] NOAEL, 1 hr = 75 ppm (0 died/2 exposed)	NOAEL	Monkey	100	0.75	0.03	*	*	*
Nasal Injury[33] 12/42 exposed to 0.2 ppm, 30 cc, single sniff	Mod	Man	1	0.002	0.002	0.002	0.002	0.002
Heinz Bodies[36] 10 min @ 90 ppm	Mild	Man	1	15	*	*	*	*
Carcinogen[34] (b) 2 ppm, 1 yr, intermittent	Mod	MiceCr&Howe model		*	*	3.5	0.85	0.15
Liver Effects[35] 0.04 ppm, 90 d, contin	NOAEL	Dogs	10	*	*	0.05	0.012	0.002
Blood Effects[35] 0.04 ppm, 90 d, contin	NOAEL	Dogs	10	*	*	0.05	0.012	0.002

* N/A for use of "concentration times time" rule.

Using the equation for the linearized multistage model shown above for HZ, MACs were calculated using a continuous lifetime exposure to 0.01 ppm for which the NRC COT calculated an upper 95% risk of 0.01. Because the model is conservative, no safety factor is used to convert animal test data to human exposure limits. Since it is not anticipated that microgravity will affect humans' sensitivity to the potential carcinogenic effects of MH, no adjustments were made for microgravity-induced physiological changes.

Liver Effects:

In continuous, 90 day low dose inhalation exposures of rats, dogs and monkeys at 0.04 pm and 0.1 ppm, dogs were the only species in which liver pathology was observed, and only at the higher dose[35]. MAC values were calculated using Haber's rule and a safety factor of 10 to the NOAEL level (0.04 ppm).

For each exposure duration, the SMAC was set equal to the lowest MAC value among the endpoints shown above. The resulting values, in units of both ppm and mg/m^3, are listed along with the most sensitive target toxicity at that exposure duration.

TABLE 4: SMACS for MH

	ppm	mg/m^3	Target Toxicity
1-h SMAC	0.002	0.004	Nasal Injury
24-h SMAC	0.002	0.004	Nasal Injury
7-d SMAC	0.002	0.004	Nasal Injury
30-d SMAC	0.002	0.004	Nasal Injury
180-d SMAC	0.002	0.004	Nasal Injury/Blood Effects/Liver Effects

*Previous 7-d SMAC was 0.04 ppm.

624

CONCLUSION

While one would intuitively expect that the toxicities of HZ and MH would be similar, the SMAC values appear to contradict this. We feel that this situation is an artifact of the quality and type of studies which have been performed to date. For MH, the SMAC values are based on the results of a single study, which showed a much higher toxicity for MH than all other studies we found. In spite of this inconsistency, the study appears to be well done and its results cannot be ignored. For HZ however, no similar test has been performed, to our knowledge. Thus, no data are available in the literature to support lowering the SMAC for HZ to a value comparable to the SMAC for MH. The old scientific refrain, "more work needs to be done", certainly applies to toxicological studies of the hydrazines.

REFERENCES

1. Vernot, E. H.; MacEwen, J. D.; Bruner, R. H.; Haun, C. C.; Kinkead, E. R.; Prentice, D. E.; Hall, A, III, Schmidt, R. E.; Eason, R. L.; Hubbard, G. B.; Young, J. T. (1985) Long-Term Inhalation Toxicity of Hydrazine, Fundamental and Applied Toxicology 5, 1050-1064.

2. Karstadt, M.; Bobal, R. (1982). Availability of Epidemiologic Data on Humans Exposed to Animal Carcinogens: 2. Chemical Uses and Production Volume., Teratog. Carcinog. Mutagen. 2, 151-168.

3. Jacobson, K. H.; Clem, J. H.; Wheelwright, H. J., Jr.; Rinehart, W. E., Mayes, N. (1955) The Acute toxicity of the Vapors of Some Methylated Hydrazine Vapors, A. M. A. Arch. Indust. Health 12: 609-616.

4. MacEwen, J. D.; McConnell, E. E.; Back, K. C. (1974). The Effects of 6-Month Chronic Low Level Inhalation Exposures to Hydrazine on Animals, AMRL-TR-74-125 Paper No. 16, pages 225-235, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

5. Smith, E. B.; Clark, D. A. (1972). Absorption of Hydrazine Through Canine Skin,Toxicol. Appl. Pharmacol., 21, 186-193.

6. van Ketel, W. G.(1964). Contact Dermatitis from a Hydrazine Derivative in a Stain Remover: Cross Sensitization to Apresoline and Isoniazide., Acta Derm.-Venerol., 44, 49-53.

7. Høvding, G. (1967). Occupational Dermatitis from Hydrazine Hydrate Used in Boiler Protection, Acta derm.-venereol. 47,293-297.

8. Evans, D. (1958). Two Cases of Hydrazine Hydrate Dermatitis Without Systemic Intoxication, British J. Industr. Med. 16, 126-127.

9. Reid, F. J. (1965). Hydrazine Poisoning, British Medical Journal,2, 1246.

10. Kulagina, N. K. (1962). The Toxicological Characteristic of Hydrazine,Toxicology of New Industrial Chemical Substances. Academy of Medical Sciences of the USSR. 4, 65-81.

11. Scales, M. D. C.; Timbrell, J. A. (1982). Studies on Hydrazine Hepatotoxicity. 1. Pathological Findings, Journal of Toxicology and Environmental Health 10, 941-953.

12. Roe, F. J. C. (1978). Hydrazine, Ann. Occup. Hyg. 21, 323-326.

13. Reid, F. J. (1965). Hydrazine Poisoning, British Medical Journal,2, 1246.

14. House, W. B. (1964). Tolerance Criteria for Continuous Inhalation Exposure to Toxic Materials, III. Effects on Animals of 90-Day Exposure to Hydrazine, Unsymmetrical Dimethylhydrazine, Decaborane, and Nitrogen Dioxide, ASD-TR-61-519 (III), Wright-Patterson Air Force Base, Ohio.

15. Kimball, R. F. (1977). The Mutagenicity of Hydrazine and Some of its Derivatives, Mutation Research 39, 111-126.

16. Herbold, B.; Buselmaier, W. (1976). Induction of Point Mutations by Different Chemical Mechanisms in the Liver Microsomal Assay, Mutation Research 40, 73-84.

17. Jain, H. K.; Shukla, P.T. (1972). Locus Specificity of Mutagens in Drosophila, Mutation Research 14, 440-442.

18. MacRae, W. D.; Stich, H. F. (1979). Induction of Sister-Chromatid Exchanges in Chinese Hamster Ovary Cells by Thiol and Hydrazine Compounds, Mutation Research 68, 351-365.

19. U. S. Department of Health and Human Services, Public Health Service, National Institute of Environmental Health Sciences (1989) Fifth Annual Report on Carcinogens: Summary, NTP 89-239: 166-168, Research Triangle Park, N. C.

20. MacEwen, J. D.; McConnell, E. E.; Back, K. C. (1974). The Effects of 6-Month Chronic Low Level Inhalation Exposures to Hydrazine on Animals, AMRL-TR-74-125 Paper No. 16, pages 225-235, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

21. Carter, V. L. Jr.; Back, K. C.; MacEwen, J. D. (1981). The Oncogenic Hazard From Chronic Inhalation of Hydrazine, AGARD Conf. Proc., AGARD-CP-309, B5/1-B5-9.

22. Steinhoff, D.; Mohr, U. (1988). The Question of Carcinogenic Effects of Hydrazine, Exp. Pathol. 33, 133-143.

23. Wald, N.; Boreham, J.; Doll, R.; Bonsall, J. (1984). Occupational Exposure to Hydrazine and Subsequent Risk of Cancer, British Journal of Industrial Medicine 41, 31-34.

24. Weatherby, J. H.; Yard, A. S. (1955). Observations on the Subacute Toxicity of Hydrazine, A.M.A. Arch. Ind. Health 11, 413-419.

25. MacEwen, J. D.; Kinkead, E. R.; Vernot, E. H.; Haun, C. C.; Hall, A. I. (1981). Chronic Inhalation Toxicity of Hydrazine: oncogenic effects, Report; ISS AFAMRL-TR-81-56; Order No. AD-A101847, 67 pages.

26. Subcommittee on Guidelines for Spacecraft Maximum Allowable Concentrations (SMACs) for Space Station Contaminants, Committee on Toxicology, Board on Environmental Studies and Toxicology, Commission on Life Sciences, National Research Council. (1990) Draft of: Guidelines For Developing Spacecraft Maximum Allowable Concentrations (SMACs) For Space Station Contaminants Pp 49-52.

27. Committee on Toxicology, Board on Toxicology and Environmental Health Hazards, Commission on Life Sciences, National Research Council, (1985) Emergency and Continuous Exposure Guidance Levels for Selected Airborne Contaminants, 5:5-21, National Academy Press, Washington, D. C.

28. Doull, J. (1989) Letter from the National Research Council's Committee on Toxicology to Colonel Thayer J. Lewis, M. C., Headquarters, U. S. Air Force, Bolling Air Force Base, Washington, D. C., dated August 8, 1989.

29. Haun, C. C.; MacEwen, J. D.; Vernot, E. H.; and Eagan, G. F., (1970) Acute Inhalation Toxicity of Monomethylhydrazine Vapor, Am. Ind. Hygiene Assoc. J., 31:667-677.

30. Comstock, C. C.; Lawson, L. H.; Greene, E. A.; Oberst, F. W. (1954). Inhalation Toxicity of Hydrazine Vapor, AMA Arch. Ind. Hyg. Occup. Med. 10, 476-490.

31. Haun, C. C.; Kinkead, E. R. (1973) Chronic Inhalation Toxicity of Hydrazine, AMRL-TR-73-125, Paper No. 25, Pages 351-363, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

32. Hoffman, E. J.; Schluter, L. A. (1976). Test Report: Olfactory Response to Monomethylhydrazine. TR-WSTF-140: NASA Johnson Space Center, White Sands Test Facility, Las Cruces, New Mexico.

33. Kinkead, E. R.; Haun, C. C.; Vernot, E. H.; Gaworski, C. L.; MacEwen, J. D.; Hall, A.; Amster, R. L.; Bruner, R. H. (1985). A chronic inhalation toxicity study on monomethylhydrazine. Springfield, VA.: National Technical Information Service; AFAMRL-TR-85-025

34. MacEwen, J. D.; Haun, C. C. (1971). Chronic exposure studies with monomethylhydrazine. Proceedings of the Second Annual Conference on Environmental Toxicology, 31 August, 1 and 2 September, 1971, NTIS # AD 751440:255-270, Wright-Patterson Air Force Base, Ohio.

35. MacEwen, J. D.; Theodore, J.; Vernot, E. H. (1970). Human Exposure to EEL Concentrations of Monomethylhydrazine, NTIS AD 727-527, 355-363, U. S. Department of Commerce, National Technical Information Service, Springfield, VA.

626

HYDRAZINE MONITORING IN SPACECRAFT

J. H. Cross*, S. W. Beck*, T. F. Limero*, and J. T. James.

Biomedical Operations and Research Branch
NASA/Johnson Space Center, Houston, TX 77058

* KRUG Life Sciences, Inc.
1290 Hercules Drive, Houston, TX 77058

ABSTRACT

Hydrazine (HZ) and monomethyl hydrazine (MMH) are highly toxic compounds used as fuels in the Orbiter main engines and in its manuvering and reaction control systems. Satellite refueling during a mission may also result in release of hydrazines. During extravehicular activities (EVAs) the potential exists for hydrazines to contaminate the suit and to be brought into the internal atmosphere inadvertently at the conclusion of an EVA. If elevated levels of these compounds are detected in the airlock, countermeasures can be instituted before the airlock is opened to the crew quarters. Because of the high toxicity of hydrazines, a very sensitive, reliable, interference-free, and real-time method of measurement is required.

Current methods for on-board measurement of hydrazine contamination in the airlock include electrochemical sensors and Draeger tubes. These methods may be adequate for measuring hydrazines at relatively high levels (ppm); however, a measurement method capable of accurately measuring concentrations in the low part per billion range ($\approx$10 ppb) is needed. A portable ion mobility spectrometer (IMS) has exhibited a low ppb detection limit for hydrazines, suggesting a promising technology for the detection of hydrazines in spacecraft air.

The Hydrazine Monitor is a modified Airborne Vapor Monitor (AVM) with a custom-built datalogger, manufactured by Graseby Ionics. This off-the-shelf IMS (the military version is known as the Chemical Agent Monitor, or CAM) was developed for the detection of chemical warfare agents on the battlefield. After early evaluations of the AVM for hydrazine measurements showed a serious interference from ammonia, the AVM was modified to measure HZ and MMH in the ppb concentration range without interference from ammonia in the low ppm range.

The Hydrazine Monitor uses a pump to draw a sample over a heated membrane. By dissolution and diffusion, vapors pass through the membrane to the analytical system. Sample molecules are ionized by reactions with a plasma of positive and negative ions, formed by ionization from a radioactive

source. Through a series of fast ion-molecular reactions, product ions are formed. Some of these ions enter the drift tube through a pulsed shutter grid. In the drift tube, the product ions drift through the cell under the influence of an electric field and arrive at the collector electrode at different times based on charge, mass, and shape. The current created at the collector electrode is amplified and a signal vs time spectrum is generated. Ammonia, HZ and MMH have characteristic drift times and spectral signatures, which permit identification and quantification of each of these compounds.

INTRODUCTION

Hydrazines are acutely-toxic and carcinogenic compounds used as rocket fuels in the military and in the space program. Monomethyl hydrazine (MMH) is used in the orbital manuvering system and reaction control system of the Space Shuttle, and hydrazine (HZ) has been selected as the fuel for Space Station Freedom. Since both compounds are also used to fuel satellites, the potential exists for hydrazines to contaminate an astronaut's space suit during an extravehicular refuelling operation and thus inadvertently be brought into the airlock at the end of the EVA. Detecting these compounds is necessary before the airlock is opened to the crew quarters. But detection is difficult. The 7-day Spacecraft Maximum Allowable Concentration (SMAC) for both HZ and MMH is 0.04 ppm.[1] Much lower limits have been proposed for new 7-day SMACs and for SMACs for longer exposures (Personal Communication, H. D. Garcia). Furthermore, the American Conference of Governmental and Industrial Hygienists (ACGIH) has recently proposed an 8-hour, time-weighted-average exposure limit of 0.01 ppm for industrial workers. Measuring such low levels quickly and reliably during operations challenges analysts and instrument makers.

The emergence of IMS as a sensitive, reliable, portable, real-time technique for determining chemical warfare agents spurred our interest in applying IMS to the measurement of hydrazines. A preliminary study of a Graseby Ionics AVM by

Dr. Gary A. Eiceman at New Mexico State University[2] was followed by a more detailed study conducted in the Toxic Vapor Detection Laboratory at Kennedy Space Center (KSC) by Ms. Rebecca Young and Dr. Eiceman.[3] They found that the AVM could detect low ppb levels of HZ and MMH, although response and recovery times were longer than desirable and ammonia interfered with the determination.

This positive study encouraged further interest at both KSC and Johnson Space Center (JSC). The interest at KSC arose from the need to monitor very low level exposure during numerous fueling operations. At JSC the issue has been the potential for exposing crews to hydrazines during EVAs. In the past, Draeger tubes and electrochemical measurements were the only viable inflight techniques available; neither method, however, possesses the sensitivity to detect hydrazines at the SMACs. Presently exposure is avoided by purely procedural controls; collecting exposure data is seen as a vital confirmation of the effectiveness of procedural controls. Both NASA centers are funding further development of IMS sensors.

A key step in the further development of IMS for HZ monitoring was removing the interference from ammonia. After Dr. Eiceman and one of us (T.F.L.) demonstrated a promising solution, the JSC Toxicology Group developed a Detailed Test Objective to conduct inflight tests (DTO 640 Hydrazine Monitor). The monitor's first flight was on STS-37 in April 1991. The major objectives were: 1. to assess instrument operation during flight, 2. to evaluate the instrument's ability to remain calibrated during flight, 3. to estimate the effect of ammonia on detecting hydrazines, and 4. to search for the presence of unanticipated compounds in the Shuttle cabin that could interfere with hydrazines' detection. The test plan included calibrations before launch and after landing of STS-37. These calibration tests were designed to collect data on the measurement range, the response times, and the limit of detection for MMH. Other tests were conducted to investigate the resolution of hydrazines from other compounds, especially ammonia, that might interfer.

This paper discusses the following results from these tests:

    1. Measurement ranges for HZ and MMH

    2. HZ and MMH calibration--linearity and stability

    3. Response and recovery times

    4. Resolution between HZ, MMH and ammonia

    5. Flight results--background interferents and other observations

METHODS

HZ and MMH Generation

The apparatus for generating ppb concentrations of hydrazines

is installed in the Toxic Vapors Detection Laboratory at KSC. A detailed description of its operation is included in reference 3, but briefly, the hydrazine permeates at a controlled rate through a plastic tube into a humidified air stream. The flow rate of the air stream is increased until the hydrazine concentration drops to the desired level. The most important fact about the generation method is that the final concentration is verified by an independent, reliable analysis. This verification at the point of delivery compensates for losses due to the reactivity and adsorptivity of hydrazines. The air stream temperature was 24-25 $^\circ$C, and the relative humidity of the zero and calibration gases was 45$\pm$5%. The nominal HZ and MMH concentrations were 40, 100, 300, 600 ppb. The lowest verified concentration available when these calibrations were performed was 9 ppb MMH. Therefore, the limit of detection for the hydrazine monitor was probed using MMH.

Theory of Ion Mobility Spectrometry

Ion mobility spectrometry separates ions by differences in the time it takes them to drift through a gas at atmospheric pressure and an applied electrostatic field[4] (Fig 1). In the instrument described in this paper sample, is introduced across a silicon rubber membrane. Provision is made for a slow airflow (not shown in the figure) to carry the sample to an ionizer, most commonly a nickel-63 radioactive foil. Through a complex process, the beta particles in the ionizer generate a reservoir of positive and negative ions known as reactant ions. The reactant ions undergo complex ion/molecule reactions with the sample molecules in the

reaction region. Compounds with high proton affinity such as the hydrazines form positive ions and those with high electron affinity such as halogenated compounds form negative product ions. The reactant ions influence the measurement range of the IMS. Fast, complete reaction of a sample with the reactant ions leads to greater sensitivity (i.e. a lower detection limit). The upper limit to the measurement range is reached at the sample concentration that depletes the reactant ions.

At the end of the reaction region, a fraction of the positive or negative product ions and excess reactant ions (of the same charge) are allowed to enter the drift cell through a pulsed shutter grid. Under a constant electric field, each ion species migrates through the drift gas to the collector at a characteristic velocity. Reversing the polarity of the shutter grid and the electric field allows ions of opposite charge to enter the drift tube. The current generated at the collector plate is converted to a voltage, amplified, measured as a function of time, and a spectrum is generated. Signal averaging increases signal-to-noise ratios. A typical IMS spectrum (bottom of Fig. 1) contains a reactant ion peak (RIP) and product ion peaks, which can number from one to many depending on conditions in the IMS. The IMS conditions are also arranged to produce suitable drift times, between 10 and 20 msec for this work.

A technique used often in IMS to improve selectivity for a particular class of compounds is to alter the ionization

R^+ = Reactant ion
A^+ = monomethyl hydrazine
B^+ = hydrazine
C^+ = ammonia

Figure 1: Principles of Ion Mobility Spectrometry

chemistry by adding a dopant chemical to the drift gas. The ions from the dopant become the reactant ions, and their reaction with the sample molecules forms ions that selectively separate from interfering compounds. The constant concentration of dopant necessary is usually maintained with permeation or diffusion tubes in the drift gas stream.

A dopant was used to eliminate the ammonia interference with hydrazines. Under a contract from the JSC Toxicology Group, Dr. Eiceman found that 5-nonanone provided resolution between MMH, HZ and ammonia without loss of sensitivity (G. A. Eiceman and T. F. Limero, manuscript in preparation). Graseby Ionics, Ltd, then built two AVMs using the nonanone dopant. Selectivity for hydrazines was high (ppb levels could be seen in the presence of ppm levels of ammonia), and sensitivity for hydrazines remained in the ppb range.

The Hydrazine Monitor

The Graseby Ionics, Ltd, AVM is a commercial, hand-held IMS (Fig. 2). The nozzle is at the right covered by a filter cap that provides clean air so the RIP can be identified during warmup. Sample drawn through the nozzle by a pump is evenly distributed across the heated membrane behind the

white nozzle to speed transfer and thereby reduce response time. The drift tube is immediately behind the membrane. Since it runs parallel with the center line of the instrument, no part of it is visible in the photograph. Behind the drift tube are sieve packs, and around it is a circuit board. The figure shows the display panel, which indicates exposure levels with 0-8 bars. In addition, a three-dot display appears when high sample concentrations deplete the RIP. The handle contains a nickel cadmium battery pack. On the bottom is a datalogger for storing spectra inflight. As shown, the connector at the back of the AVM sends IMS spectra to the datalogger. The spectra can also be sent to a computer for viewing.

RESULTS AND DISCUSSION

Measurement Ranges

The upper measurement limit for MMH was close to 600 ppb, since the bar graph read 8 bars, the maximum, and the three dot display lit.

The lower limit of detection for MMH was probed in an experiment. MMH can be detected at 9 ppb. The MMH peak had a drift time of 15.4 msec (Fig. 3). The height of the peak

Figure 2: Prototype Hydrazine Monitor with Data Logger

was measured by comparing the voltage reading at the top of the peak with the voltage reading at the same drift time after the MMH was allowed to clear from the AVM. The peak height was 8 mV, small but detectable. Since the MMH peak was clearly discernable, one can speculate that an even smaller concentration could be detected by using more sophisticated signal-to-noise averaging techniques. The peak at 16.3 msec is an impurity in the diluent air; the impurity served as a reference that confirmed the presence of the MMH peak. Note that this experiment shows that at least one of the hydrazines can be detected below the proposed 10 ppb exposure limit.

Hydrazine produced a larger response from the AVM than did MMH. This observation is consistent with HZ's greater polarity.

The upper measurement limit for HZ was above 600 ppb (the maximum concentration tested). Although the spectrum for 600 ppb showed a small RIP, the three-dot display indicating RIP depletion was not lit on the display, and therefore the measurement limit was thought to be slightly higher. Considering the size of the RIP at 600 ppb, one can speculate

that the upper measurement limit would be about 700 ppb. The bar graph display showed 8 bars (maximum) for 600 ppb.

The lowest HZ concentration we have tested, 40 ppb, gave a strong signal, and therefore the limit of detection is lower. The limit for HZ was inferred from the MMH limit. The peak height for MMH at 9 ppb was 8 mV. For one 40 ppb HZ replicate, the peak height was 80 mV. A HZ peak 8 mV high would thus represent 4 ppb HZ.

Calibration--Linearity and Stability

Calibration was performed over a range of greater than one order of magnitude. Since this was the first time the monitor had been calibrated against known concentrations of HZ and MMH, we focused on examining the relationships between concentration and response. The data were examined several ways; this discussion will emphasize linear relationships across the entire range.

The most nearly linear response to HZ came from the percentage of the HZ peak areas to the total peak areas in the spectrum. As the HZ concentration increased, the RIP

GAL Monitor 08:52:34 22/03/91 Loading From c:\data\mmh9068.fnh
Comment - AVM 1 CALIBRATION: 9.0 ppb MMH

Cursor   0, 5000uS, 466mV   Window 5000- 5000us   Peak 5033uS,   0mV
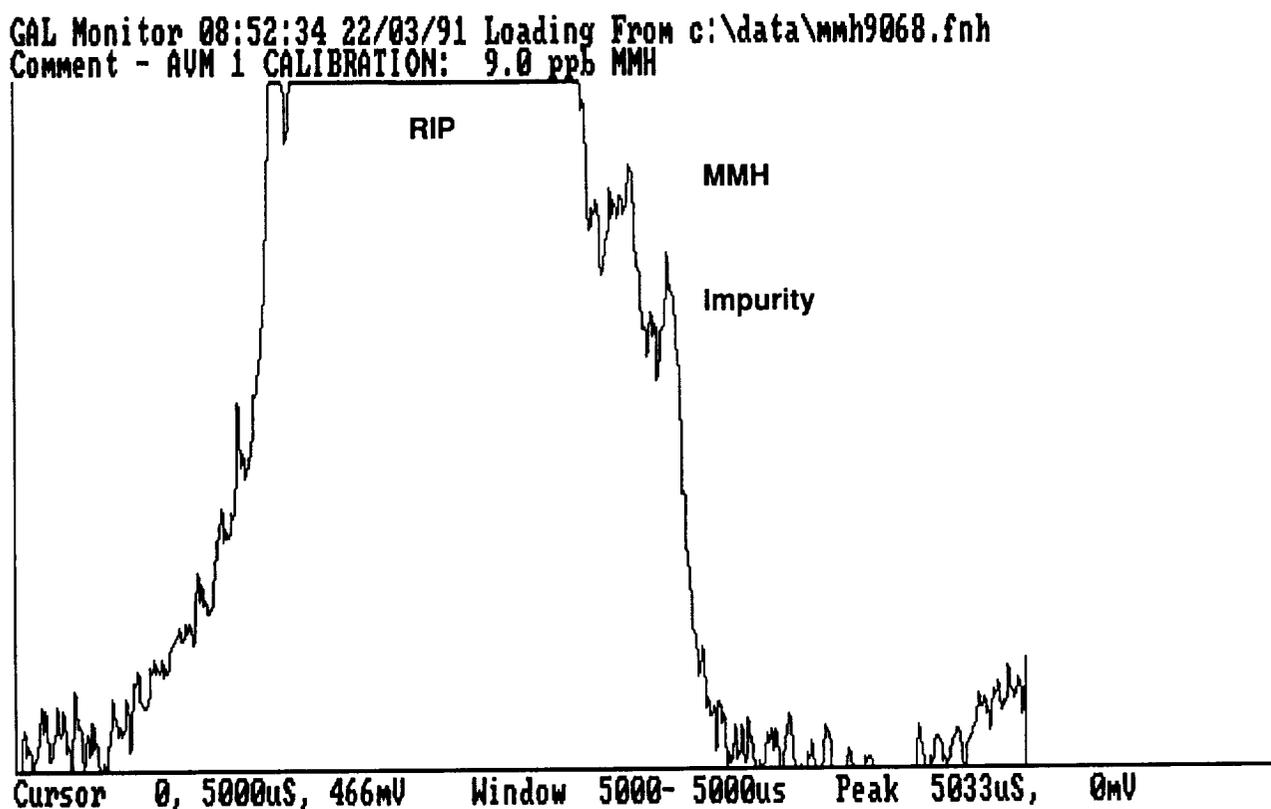
Figure 3: IMS Spectrum of 9.0 ppb MMH. The signal from 1024 scans has been averaged.

631
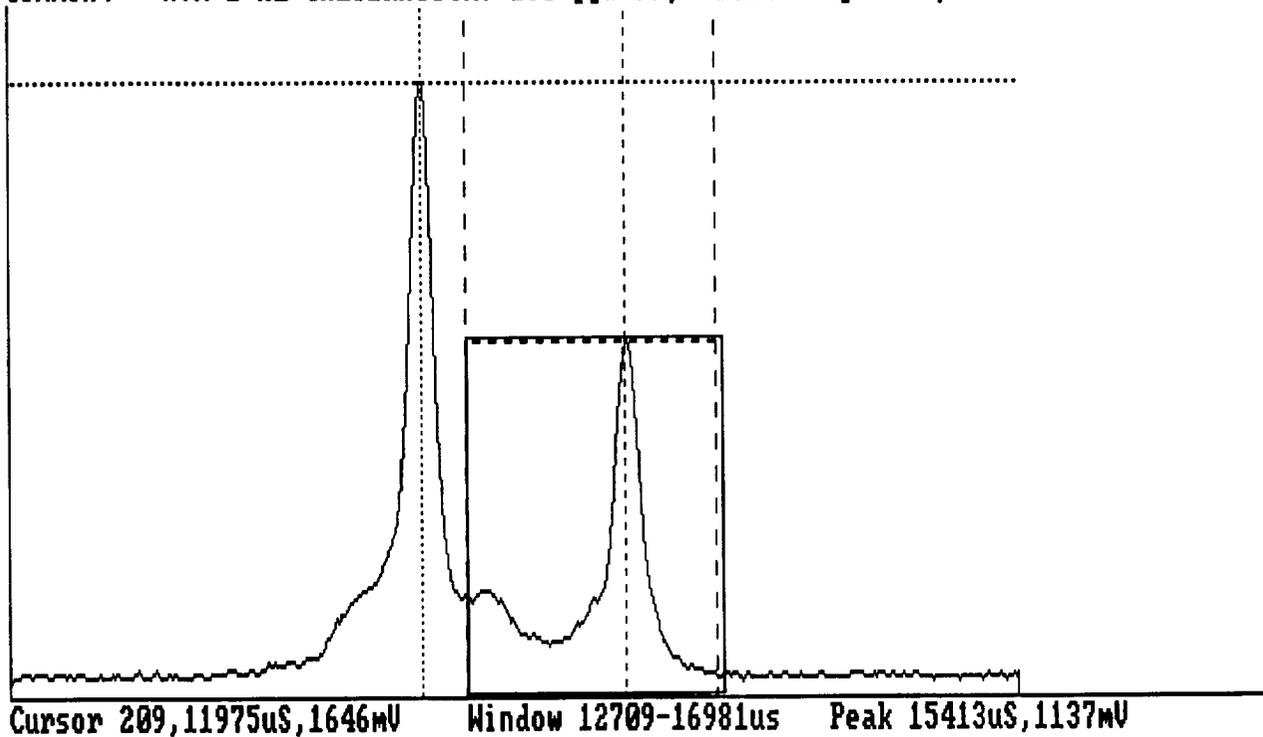
Figure 4: Spectrum of 300 ppb HZ
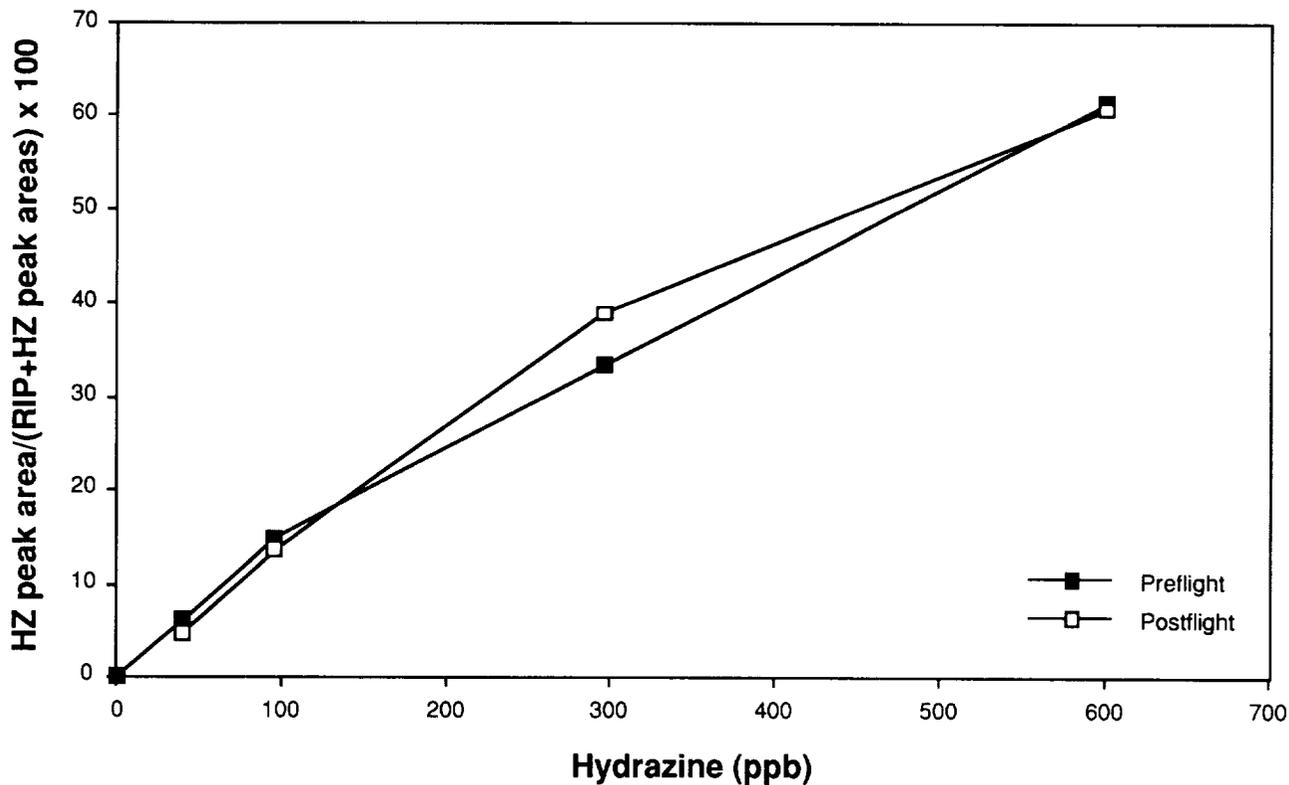
Figure 5: AVM Calibration Curves for HZ Before and After STS-37
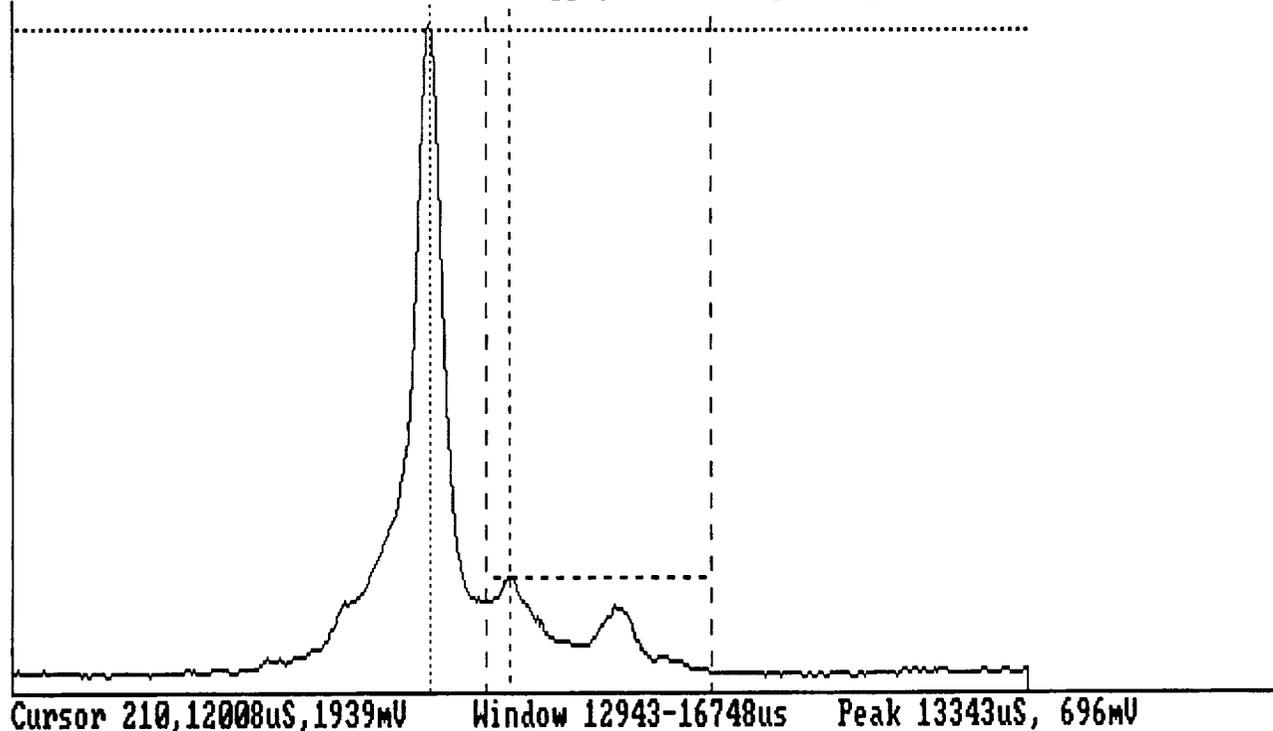
Figure 6: Spectrum of 100 ppb MMH

decreased and two HZ peaks grew proportionately (Fig 4&5). The total area under the peaks appeared to remain constant up to 600 ppb. The preflight HZ response (as defined above) was linear across the range. In the postflight HZ calibration, the peak area method was also linear (Fig. 5). In this postflight calibration, the linearity of the peak areas was not clearly superior to the linearity determined from the height or the area of the larger HZ peak (the rightmost peak in the spectrum).

Calibrating with peak area worked less well with MMH than with HZ. The total area under the RIP and the two MMH peaks decreased as the concentration of MMH increased, and the curve flattened at higher concentrations.

The most linear response to MMH came from the peak height of the MMH peak at a drift time of 13.3 msec (Fig. 6&7. The vertical dashed line on the smaller peak shows the 13.3 msec peak.). Since at low concentrations, this peak was a shoulder on the RIP, the peak height had to be measured above the "instrument baseline", the baseline at a drift time well away from any peaks, to avoid confusion caused by shifts in the RIP shape. Because the MMH peak at 13.3 msec was a shoulder peak, it was harder to detect at lower concentrations. In contrast, the MMH peak at 15.4 msec became nonlinear at

moderate concentrations, but was easier to detect at lower ones.

The correspondence between calibrations before and after flight was remarkably good. Note also that in Fig. 7 the 9 ppb MMH point falls on the curves. For both HZ and MMH, the voltage readings corresponding to particular concentrations were nearly the same for the preflight and postflight calibrations (for MMH, see Fig. 7). This stability indicates that an IMS can undergo the rigors of spaceflight without requirements for on-orbit adjustments. In terms of AVM performance, the calibration stability suggests that the drift tube, the component where the analytical determination takes place, gives stable responses over long times and with active AVM usage.

Response and Recovery Times

Although the AVM detected 40 ppb concentrations of HZ and MMH in 2 min, stable response times for both HZ and MMH were lengthy, from 15 to 20 minutes. The long response time for hydrazine was partially mitigated by the great sensitivity of the AVM; after 2 minutes, the hydrazine peak was measurable and its size could be extrapolated to predict the final concentration. While 9 ppb MMH was identifiable at 10

633
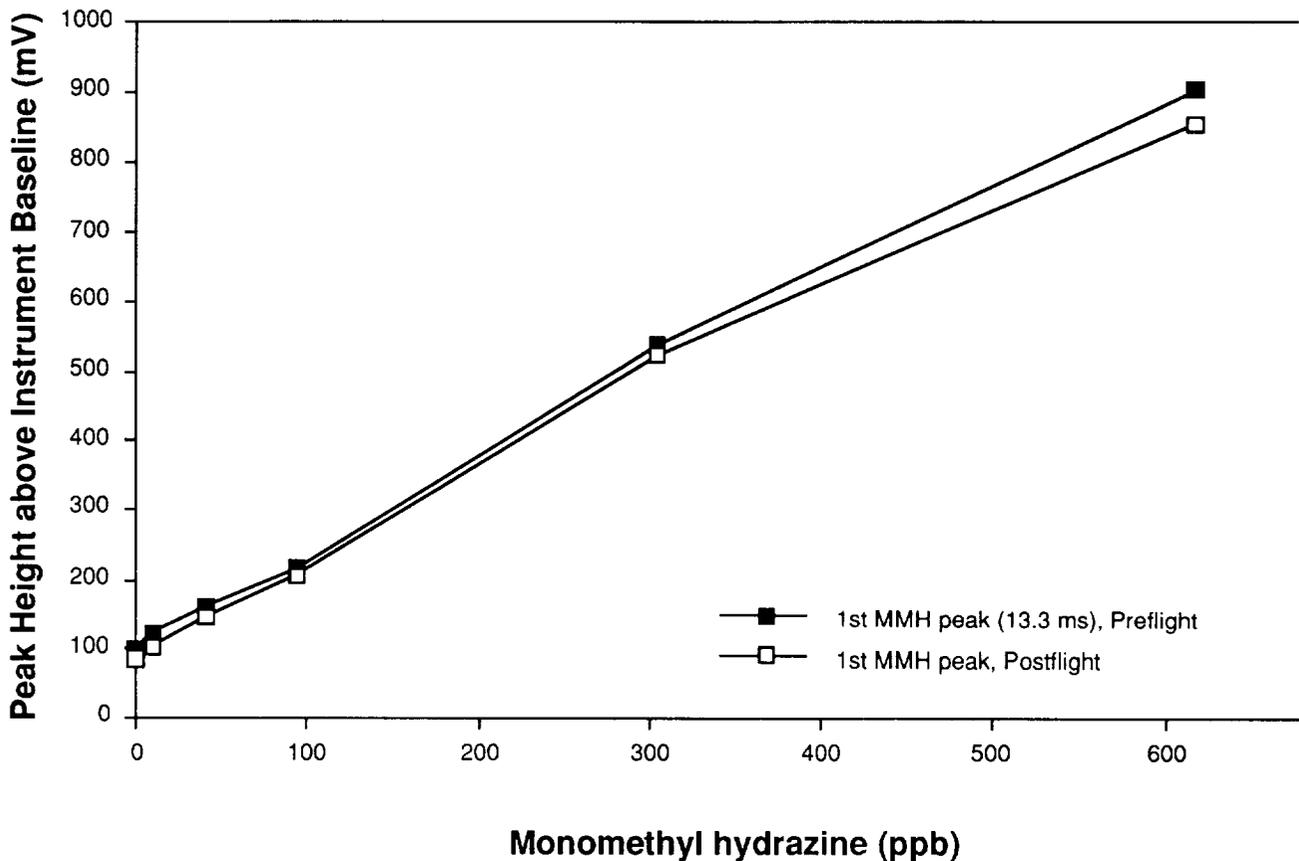
Figure 7: AVM Calibration Curves for MMH Before and After STS-37

min, quantification of 9 ppb required a 60 min exposure of the AVM.

At lower concentrations, 40-100 ppb, the recovery times for HZ and MMH were 10-15 min. At higher concentration, the recovery times for HZ (>40 min) were longer than for MMH (≈30min), which was consistent with HZ's greater polarity. While long recovery times are undesirable, they do indicate that decomposition of these very reactive rocket fuels was not the primary cause of long response times. Decomposition of the sample is regarded as a problem in the analysis of hydrazines, and this indication that its importance is low encourages further development.

Considering the excellent drift tube performance discussed above, one can speculate that the long response and recovery times arise from another component of the AVM. The high polarity of hydrazines suggests that they may be adsorbing strongly in the nozzle or the membrane.

Resolution

The results described above were obtained with unmixed hydrazines. Identification and quantification were straightforward based on the drift times of the two peaks for

each. We have also found that ammonia was well separated; its drift time, 16.4 msec, was longer than either HZ or MMH. Ammonia could be present at perhaps 8 ppm before its peak would broaden to obscure hydrazine. Figure 8 shows a spectrum of all three compounds. The ammonia concentration was roughly 4 ppm, and the peak heights of HZ and MMH show them to be roughly 100 ppb. Distinguishing between HZ and MMH would be relatively easy, because each has two peaks. Quantification and identification of mixtures of low concentrations of both HZ and MMH may require be more difficult, since the one peak in the spectrum of HZ and MMH have the same drift time.

Flight Results

Data were collected at three locations at five times for a total of 15 data sets. The sampling locations were the airlock, the galley and the waste control system (WCS) at the following times:

> Before prebreathe (Cabin pressure=14.7 psia, nominal)
> Before contingency EVA (10.2 psia)
> After contingency EVA (10.2 psia)
> Before scheduled EVA (10.2 psia)
> After scheduled EVA (10.2 psia)

634

```
GAL Monitor 13:26:54 28/06/91 Loading From b:\spectra\avmlhi83.all
Comment - AVMH 1001 HZ, MMH, NH3. 2.0L/MIN, 50C
```

```
Cursor 435,16518uS,1104mV      Window 8307-17486us    Peak 12112uS,1411mV
```
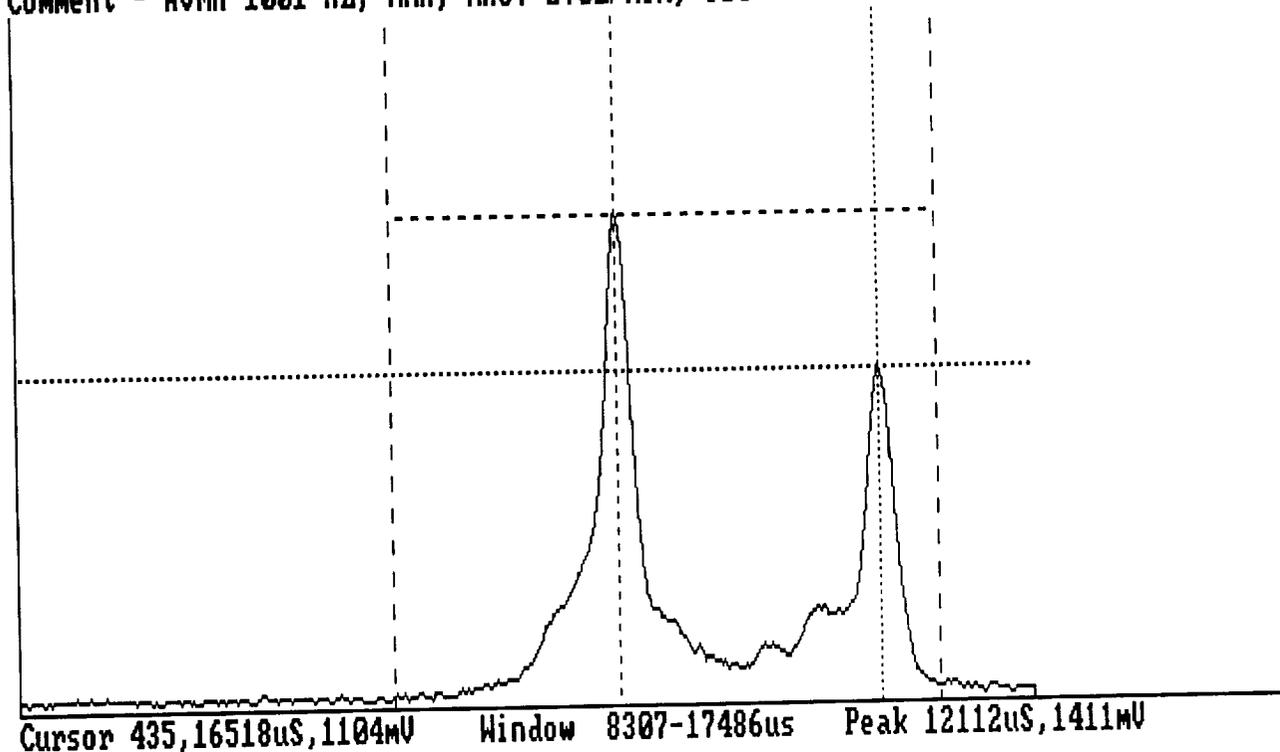
Figure 8: Spectrum of Approx. 100 ppb each HZ and MMH and Approx. 4 ppm of Ammonia

No interfering compounds were detected at any location. Small peaks did appear before the RIP in two data sets, and ammonia peaks just above the noise level were tentatively identified in several data sets, but no significant peaks appeared for hydrazines or any other compound.

Cabin pressure changes caused the RIP to shift position. In a postflight experiment, drift time was linear and directly proportional to the pressure; a pressure change of one psia changed the RIP drift time by 0.8 msec. Pressure changes affect all ions equally[4]. Thus, the resolution between ions is unaffected. In other words, the proportional separation between ions remains the same during pressure changes. The AVM uses proportional counting to locate the hydrazines relative to the RIP, which it identifies as the largest peak in the spectrum during warmup.

Summary and Conclusions

The hydrazine monitor was stable and sensitive with HZ producing a large response at 40 ppb and MMH detectable at 9 ppb. The instrument maintained linear calibration from <10 ppb to >600 ppb for over a month and throughout a spaceflight. Replication between calibrations before and after

flight was excellent. Resolution between ammonia at >4 ppm and HZ and MMH was at least 20 to 1. Although response times were longer than desired, probably because the polar hydrazines adsorbed in the inlet, both compounds could be detected in about two minutes. The instrument worked well during flight and at reduced pressures. Compounds in the Shuttle cabin atmosphere that might interfere with the determination of hydrazines are either absent or present in low concentrations. Finally, the performance of the AVM on STS-37 suggests that IMS technology is suitable for hydrazine detection in spaceflight.

Future Work

The JSC Toxicology Group is currently identifying the most important instrument modifications to be made before the STS-49 mission in April 1992, which is the next mission with a scheduled EVA. Modifications to the nozzle and membrane are likely to reduce the response and recovery times. Data reduction algorithms specfic for detecting and quantifying hydrazines will be developed. After these improvements are incorporated into the AVM and the datalogger, another flight test will be conducted to insure the absence of interfering compounds in the spacecraft environment.

635

REFERENCES

1. Coleman, M. E. and James, J. T., "Spacecraft Maximum Allowable Concentrations for Airborne Contaminants," Technical Publication No. 20584, NASA/Johnson Space Center, Houston, Texas, May 1990.

2. Eiceman, G. A. and Leasure, C. S., "Continuous Detection of Hydrazine and Monomethylhydrazine Using Ion Mobility Spectrometry," Anal. Chem., Washington, D. C. 57(9), Aug 1985, 1890-1894.

3. Young, R., "A Report on Test Results on Ion Mobility Spectrometer Detection of Hydrazines," DL-ESS-24, NASA/Kennedy Space Center, Titusville, Florida, June 1989.

4. Hill, H. H., Siems, W. F., St. Louis, R. H., and McMinn, D. G., "Ion Mobility Spectrometry," Anal. Chem., Washington, D. C., 62(23), Dec. 1990, 1201A-1209A.

636

|

COMPARISON OF DERMAL AND INHALATION ROUTES OF ENTRY FOR ORGANIC CHEMICALS

Gary W. Jepson
Toxic Hazards Division
Det 1 AL/OETA
WPAFB, OH 45433

James N. McDougal
Toxic Hazards Division
Det 1 AL/OET
WPAFB, OH 45433

Harvey J. Clewell III
Hazardous Materials Safety
WPAFB, OH 45433

ABSTRACT

The quantitative comparison of the chemical concentration inside the body as the result of a dermal exposure versus an inhalation exposure is useful for assessing human health risk and deciding on an appropriate protective posture. In order to describe the relationship between dermal and inhalation routes of exposure, a variety of organic chemicals were evaluated. The types of chemicals chosen for the study were halogenated hydrocarbons, aromatic compounds, non-polar hydrocarbons and inhalation anesthetics. Both dermal and inhalation exposures were conducted in rats and the chemicals were in the form of vapors. Prior to the dermal exposure, rat fur was closely clipped and during the exposure rats were provided fresh breathing air through latex masks. Blood samples were taken during 4-hour exposures and analyzed for the chemical of interest. A physiologically based pharmacokinetic model was used to predict permeability constants (cm/hr) consistent with the observed blood concentrations of chemical. The ratio of dermal exposure to inhalation exposure required to achieve the same internal dose of chemical was calculated for each test chemical. The calculated ratio in humans ranged from 18 for styrene to 1180 for isoflurane. This methodology can be used to estimate the dermal exposure required to reach the internal dose achieved by a specific inhalation exposure. Such extrapolation is important since allowable exposure standards are often set for inhalation exposures, but, occupational exposures may be dermal.

INTRODUCTION

Dermal penetration may be the primary mode of entry for occupational and environmental chemicals (1). This is particularly true in occupational settings where workers are provided with respiratory protection but have skin exposed to chemical vapors. In situations where the chemical vapor concentrations are high enough and exposure is long enough, the amount of chemical absorbed into the body through the skin can produce a chemical body burden up to or exceeding the levels achieved by an inhalation exposure at the threshold limit value (TLV). Since the skin can function as a significant route of entry for chemicals which have regulated inhalation exposure standards, it is important to quantitatively assess sthe contribution of dermal penetration to the internal dose of chemical resulting from an exposure (2). In order to compare the internal chemical dose following inhalation and the internal dose following dermal exposure, physiologically based pharmacokinetic (PBPK) models are used to account for the uptake, biodistribution, metabolism and elimination of chemicals. The models are built by connecting a series of physiological compartments which have descriptions of blood flow, tissue volume and if appropriate, metabolism (3). Such a description allows for tracking the chemical mass balance as well as the kinetics of the physiological processes impacting the fate of the chemical. When combined with appropriate laboratory data, the PBPK model provides a useful tool for comparing the amount of chemical absorbed following inhalation or dermal exposure.

The total amount of chemical absorbed is the primary piece of information required to make the quantitative comparison between different routes of chemical exposure and allows for derivation of a dermal exposure that would yield a chemical body burden equivalent to that produced by a selected inhalation exposure.

RESULTS

Chemical concentration profiles in rat blood during dermal exposure to chemical vapors combined with use of a physiologically based pharmacokinetic model provided the information needed to calculate dermal permeability constants for the chemicals used in the study (Table 1).

Table 1. Dermal Vapor Penetration

Chemical	Permeability(cm/hr)
Styrene	1.75
Dibromomethane	1.32
Bromochloromethane	0.79
m-xylene	0.72
Tolene	0.72
Perchloroethylene	0.67
Methylene chloride	0.28
Benzene	0.15
Halothane	0.05
Hexane	0.03
Isofurane	0.03

Similar chemical concentration profiles in rat blood during an inhalation exposure and use of a PBPK model allowed for calculation of the internal dose that resulted from the exposure. The ratios of dermal to inhalation exposure concentration which yield identical internal doses or chemical body burden are shown in Table 2.

Table 2. Extapolation Ratios

Chemical	Ratio (Dermal/Inhalation)	
	Rat	Human
Styrene	10	18
Dibromomethane	14	23
Bromochloromethane	22	39
m-xylene	25	42
Tolene	24	43
Perchloroethylene	26	46
Methylene chloride	62	110
Benzene	113	202
Halothane	339	682
Hexane	408	990
Isoflurane	448	1180

CONCLUSION

Dermal exposure to chemicals of occupational or environmental interest can result in an internal dose or chemical body burden which may be unacceptable based on the internal dose received following an inhalation exposure for which exposure limits exist. The use of a physiologically based pharmacokinetic model combined with the generation of kinetic data during dermal and inhalation chemical exposures, allows for the calculation of an extrapolation ratio. The extrapolation ratio is the ratio of the dermal and inhalation exposure concentrations required to produce the same internal dose of chemical. Therefore, the extrapolation ratio can be used to assign a dermal TLV for chemical vapor exposure if an inhalation TLV exists.

REFERENCES

1. J.N. McDougal, G.W. Jepson, H.J. Clewell, M.G. Gargas and M.E. Andersen, Fundam.Appl.Toxicol. 14:299-308, 1989.

2. J.N. McDougal, G.W. Jepson, H.J. Clewell, M.G. MacNaughton and M.E. Andersen, Toxicol.Appl.Pharmacol. 85:286-294, 1986.

3. J.C. Ramsey and M.E. Andersen, Toxicol.Appl.Pharmacol. 73:159-175, 1984.

OCCUPATIONAL SAFETY CONSIDERATIONS WITH HYDRAZINE

Lt Col Harvey J. Clewell III,
Toxic Hazards Division
Harry G. Armstrong Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

Major James N. McDougal, PhD
Marilyn E. George
Melvin E. Andersen, PhD

ABSTRACT

Hydrazine is a reducing agent that is most commonly used as a
propellant and as an oxygen scavenger in boilers. Hydrazine is extremely
irritating and has been demonstrated to produce both acute and chronic
toxicity. As a result, the established permissible inhalation exposure
limits are very low, and respiratory protection is required whenever
vapors are present. Liquid hydrazine penetrates the skin and produces a
chemical burn; therefore, some protective measures must also be taken to
protect the skin from liquid contact. Often, however, a cumbersome,
whole-body protective suit is worn to protect against skin contact with
vapor as well. To what extent it is actually necessary to protect skin
from vapor penetration had not previously been demonstrated. In an
attempt to answer this question, we conducted a study with rats to compare
the dermal penetration of hydrazine vapor with inhalation.
Pharmacokinetic modeling was used to compare body burdens resulting from
these different routes of exposure. The analysis concluded that the
vapor concentration during a skin-only exposure would have to be at least
200 times higher than that during inhalation to achieve the same body
burden. This type of estimation illustrates the use of predictive
toxicology in occupational exposures.

I

Session H8: SPACEFLIGHT EXPERIMENTS: SURVIVING THE PROCESS

640

EXPERIMENTS TO BE FLOWN IN AN EARTH ORBITING LABORATORY: THE U.S. EXPERIMENTS ON THE FIRST INTERNATIONAL MICROGRAVITY LABORATORY, FROM CONCEPT TO FLIGHT

C. M. Winget, P. X. Callahan, R. L. Schaefer,* and J. J. Lashbrook

Space Life Sciences Payloads Office
NASA Ames Research Center
Moffett Field, California 94035-1000

*Lockheed Engineering & Sciences Company
Ames Research Center
Moffett Field, California 94035-1000

ABSTRACT

A space life sciences flight experiment requires careful long-term planning. A key role of the Space Life Sciences Payloads Office is to guide the orderly development of a proposed experiment originating in a ground-based laboratory to be flown in an Earth-orbiting laboratory. The first International Microgravity Laboratory, to be launched in 1991, provides an example of the experiment development requirements necessary to ensure a maximum science return. All life science experiments defined for spaceflight have gone through a rigorous and competitive evaluation process: a peer review for scientific quality, a program review for relevancy, and an engineering review for feasibility. Less than 10% of all proposals to do life science experiments in space are accepted for definition as candidates for flight. The candidate experiments have limited options for a flight assignment (e.g., spacelab, secondary payload, and international cooperative flights). The flight assignment is based primarily on the experiment weight, orbital requirements, services (i.e., power, cooling, etc.), and crew time requirements. To maintain the science fidelity of the experiment, an experiment requirements document (ERD) is prepared by NASA in conjunction with the Principal Investigator (PI). This ERD is then used, again in conjunction with NASA and the PI, to define the hardware requirements and generate a hardware requirements document. A phased set of reviews (e.g., preliminary requirements review, preliminary design review) is held, culminating in a critical design review of the experiment when the hardware and experimental design is approximately 90% complete. The compatibility of the science with the hardware is then further evaluated with a biocompatibility test. The final science evaluation is the Experiment Verification Test, in which flight hardware is used to perform the experiment in a simulation of the flight. All the approximately 60 steps involved in placing a life science experiment in space are coordinate with the PI by a NASA research scientist, the Payload Scientist.

ACRONYMS

AO	Announcement of Opportunity
ARC	Ames Research Center
CDR	Critical Design Review
DPM	Deputy (FPI) Payload Manager
DPS	Deputy (FPI) Payload Scientist
ERD	Experiment Requirements Document
ESS	Experiment Support Scientist
EUE	Experiment Unique Equipment
EUH	Experiment Unique Hardware
EVT	Experiment Verification Test
FPI	Flight Payload Integration
IERD	Integrated Experiment Requirements Document
IRR	Integration Readiness Review
IML-1	First International Microgravity Laboratory
KSC	Kennedy Space Center
LSLE	Life Sciences Laboratory Equipment
MST	Mission Sequence Test
NASA	National Aeronautics and Space Administration
NRA	NASA Research Announcement
PDR	Preliminary Design Review
PI	Principal Investigator
PM	Payload Manager
PMM	Payload Mission Manager
POCC	Payload Operations Control Center
PS	Payload Scientist
PRR	Preliminary Requirements Review
QA	Quality Assurance
SLSPO	Space Life Sciences Payloads Office
STS	Space Transportation System
TMA	Test Monitoring Area

INTRODUCTION

This report presents the current life cycle of NASA Ames Research Center (ARC)–managed flight experiments. The report has two main purposes: The first is to bring to the attention of biologists, and in particular cell and plant biologists, some of the requirements for flying a life science experiment in space. The second is to introduce the subject to biologists embarking on studies in the field and to delineate some of the specific requirements that will be encountered by an ARC-managed microgravity experiment. This report is not intended to be an exhaustive encyclopedia of all techniques used to prepare an experiment to evaluate the effect of microgravity on plant and animal cells. However, many of the requirements are the same for all biological systems and for other NASA centers. A detailed presentation can be found in *Principal Investigator Handbook*, ARC Rev. 12/16/90.

This document emphasizes the Principal Investigator's (PI) involvement in the activities required for successful completion of major reviews. The PI support required for activities other than these reviews is also discussed, as are the interactions between ARC and the PI that will be required as problems or questions arise throughout experiment and payload development. It is impossible to predict the extent of this activity because it varies according to the complexity of the experiment and the flight experience of the PI.

THE NEED FOR MICROGRAVITY RESEARCH

Microgravity research is a high priority for most environmental biologists because the Earth's gravitational pull is an ubiquitous environmental factor. All organisms spend their entire existence experiencing an acceleration of 9.8 m/s^2. This force acts not only on entire organisms and their organs, but also on the free-moving and sedimentable structures within their cells. Every movement and action involves a reaction against gravity. To remain erect requires an expenditure of energy, in the form of either muscular activity or energy-intensive building of stiff members (bones and wood chitin are good examples). To move requires orientation and an expenditure of energy against gravitational force. All organisms developed in this constant field over eons, and their entire structure and function are probably strongly influenced by adaptation to this force.

The absence of this force, as in spaceflight, causes environmental biologists to ask questions. What must the release of gravity mean to the functioning of the organism? What will be the reactions of organisms when movement is free; when up and down are gone; and when weight, sedimentation, and convection disappear and the energy used in reacting against them can be put to other uses? What kinds of intra- and intercellular and organ reorientation (e.g., molecular, physiological, and morphological) will take place in a gravity-free environment? It is to questions like these that biologists are able to turn their minds now that manned and unmanned earth-orbiting research facilities have become available. To meet NASA's life sciences goals, questions like these must be answered, in part to gain greater understanding of basic biological processes, and in part to aid in the realization of permanent manned orbiting facilities and interplanetary spaceflight.

BACKGROUND

Since the early 1960s biological experiments have formed a small but significant proportion of the payloads on orbiting space vehicles, and a wide variety of plants, animals, microorganisms, and cell and tissue cultures have been carried. In 1962 the Soviet Satellite 2 carried, for example, mice, guinea pigs, and human and rabbit skin. In 1967 Biosatellite II was launched by NASA with 13 selected general biology and radiation experiments (e.g., amoeba, *Tradescantia*, *Neurospora*, wheat, bell pepper, and frog eggs). Since that time, experiments on numerous Soviet and U.S. flights have added to our knowledge and peaked our interest in understanding the microgravity phenomenon. This knowledge, together with knowledge gained from ground-based research, is continually used in NASA's Space Life Sciences research planning process (Figure 1).

EXPERIMENT LIFE CYCLE

Experiments start with an initial Proposal Evaluation and progress through Experiment Definition, Experiment Development, Flight Hardware Development, Payload Integration, and Flight phases. The flow of these activities is shown in Figure 2. Unlike most ground-based biological research, space-biology experiments have lead times of years, rather than days or weeks (see Figure 3). The complexity of the development process and of the skills needed requires that an experiment team be established.

The Experiment Team

In space research, the individual researcher becomes part of a large organizational network (hundreds to thousands of people, depending on the mission), with all that it entails in terms of project management, difficulties with information flow, restricted freedom of action, and restricted flight opportunities. This report stresses that a team approach is needed to ensure a successful microgravity experiment. The goal of this team (which may, for example, consist of the PIs, ARC, and the Payload Mission Manager [PMM]) is to maximize the scientific return while minimizing the time, effort, and funds required to define, develop, and implement the experiment for space. Consequently, this team will represent the experiment to science, engineering, and operations personnel at ARC, and to NASA Headquarters, the Mission Management organization, and other center organizations involved in the flight opportunity.

Proposal Evaluation

Experiment proposals for U.S. spaceflights come from many sources. In life sciences, there are four principal sources of investigations: proposals submitted in response to a NASA Announcement of Opportunity (AO) or NASA Research Announcement (NRA), unsolicited proposals, agreements of various types made by NASA, and studies prepared in response to NASA's critical medical or technological needs. To be defined for spaceflight, all life science experiments must go through a rigorous, competitive evaluation process that includes an external peer review for scientific quality, a

program review for relevancy, and an engineering and cost review for feasibility.

Experiment Definition

Less than 10% of all proposals to do life science experiments in space are accepted for definition as candidates for flight. The preliminary experiment selection is based on the scientific merit of proposals and an initial assessment of feasibility, as indicated above. The accepted experiments are then further defined and their feasibility for spaceflight carefully evaluated against options for a flight assignment (e.g., dedicated Space-lab, secondary payload, and international cooperative flights); this is the Experiment Definition phase.

During Experiment Definition, the PI works with a member of the Payload Scientist's team. The Payload Scientist (a NASA research scientist) coordinates, with the PI, all the approximately 60 steps involved in placing a life science experiment in space; the team member involved in Experiment Definition is the Experiment Support Scientist (ESS). In this phase, the ESS works with the PI to define the experiment; this becomes an iterative process between the developing center (in this case, ARC) and the PI. The PI and the ESS establish an in-depth understanding of the experiment, which leads to the refinement of science objectives and approaches within the confines of the Space Transportation System (STS), and the evolution of resource requirements.

The Experiment Definition phase results in (1) an agreement about possible approaches to performing the experiment in space, and (2) a preliminary agreement about experiment requirements, including the number of specimens, the need for existing Life Sciences Laboratory Equipment (LSLE) (e.g., animal holding units, refrigerator, freezer, centrifuge), and the need for Experiment Unique Equipment (EUE) (unique flight equipment to be built for a specific experiment). The second agreement results in a list of tentative Spacelab resource requirements, such as experiment weight, power, and cooling, and crew time requirements. Because the Shuttle is a shared system, users are in competition with one another and with other Orbiter operations for all available resources, from mass capacity to crew time. Although a substantial level of resources is available to meet essential needs, the PI is encouraged to distinguish between requirements and desires and to use prudence in establishing the resource requirements of his/her experiment.

The Experiment Definition phase concludes with an Experiment Requirements Document (ERD) and an experiment Preliminary Requirements Review (PRR). A pool of defined candidate experiments is then formally selected by NASA Headquarters; from this pool, experiments are selected for development.

Experiment Development

During Experiment Definition or at PRR, supporting studies necessary for the development and implementation of the experiment are agreed upon by NASA and the PI. Supporting studies are generally used to demonstrate the efficacy of an approach or hardware in performing the experiment. These studies are initiated by the PI during Experiment Development (Figure 3), and they influence the design of the experiment and of Experiment Unique Hardware (EUH).

Further refinement of the experiment requirements continues until the experiment Preliminary Design Review (PDR), at which time the full ERD is baselined (no further changes will be made without a formal review). Initial design approaches are presented at this PDR and project concurrence/direction is obtained to continue design approaches or initiate new approaches. When the design is 90% or more complete, the experiment Critical Design Review (CDR) is held, at which changes ("deltas") in science requirements are presented and the experimental design is baselined.

Flight Hardware Development

During the Experiment Development effort, prototype hard-ware is developed and tested for function, support of science objectives, and biocompatibility; this leads to a full experiment test and an SLSPO acceptance review. Upon acceptance, an experiment is ready for incorporation into a payload and for development of flight hardware. Under certain circumstances, an experiment may be assigned to a payload prior to experiment acceptance. Regardless of path, flight hardware is developed in the same manner as the experiment—in stages, with a similar review cycle (PRR, PDR, CDR, acceptance).

Payload Integration

The payload development approach at ARC comprises four areas of activity: payload selection, hardware/data system development, payload development, and payload integration. Mission development at the Mission Management Center proceeds along a similar path with a similar set of reviews. There is a sequential flow of activities in each area. Activities in some areas may depend on the completion of activities in another area. For example, the experiment CDRs in the experiment development stage must precede the payload PDR in the payload development stage, which in turn must generally precede the mission PRR (see Figure 3).

To define the payload experiment requirements, resource, interface, and feasibility information that was gathered during the experiment definition and development processes is combined with information provided by the PIs and the ESSs for other experiments. This results in an Integrated Experiment Requirements Document (IERD). The series of reviews for payloads and missions reflect the developing nature of the experiments, and the experiments become more refined as the cycle progresses. Included in each review are a description of the experiment science, a list of experiment requirements, a Safety Compliance Document, and schedules.

Once the payload is selected and developed, the Experiment Verification Test (EVT) is conducted. This is a final science evaluation in which flight hardware is used to perform the experiment in a simulation of the flight. Upon completion of this test, the experiments and payload are approved by NASA

Headquarters, and the payload is integrated into the Spacelab and Shuttle in preparation for flight.

Flight

This activity includes the preparation of flight specimens, the loading of the specimens into the flight hardware, and the loading of the hardware into the Shuttle. After launch, the flight is monitored at the Flight Support Facility at Kennedy Space Center (KSC) (where ground controls may be run simultaneously with the flight experiment) and the Payload Operations Control Center (POCC) by the PI and the Experiment Team. Recovery of specimens after flight may be at KSC or at Dryden Flight Research Facility (depending on the constraints of the Mission) by other members of the Experiment Team. The specimens are then given to the PI for post-flight processing, which culminates in a PI final report. The results of this experiment are then used to further refine NASA's program objectives, thus completing the cycle (Figure 1).

Example

The first International Microgravity Laboratory (IML-1), to be launched in January 1992, provides an example of the experi-

ment development requirements necessary to ensure a maximum science return. The Space Life Sciences Payloads Office (SLSPO) at ARC developed five experiments that will fly on IML-1. Two of the experiments are Dr. Allan H. Brown's "Gravitropic Response of Plants in the Absence of a Complicating g-Force" (GTHRES) and Dr. David G. Heathcote's "Post Illumination Onset of Nutation at Zero G" (FOTRAN). Both experiments will fly in the Gravitational Plant Physiology Facility, a suite of hardware that fits into a Spacelab double rack.

The other three experiments supported by SLSPO will be flown in the European Space Agency's Biorack facility. The experiments are Dr. Gregory A. Nelson's "Genetic and Molecular Dosimetry of HZE Radiation" (US-1), Dr. Carlo V. Bruschi's "Microgravitational Effects on Chromosome Behavior of Yeast" (US-2), and Dr. Pauline Jackie Duke's "Chondrogenesis in Micromass Cultures of Mouse Limb-Bud Mesenchyme Exposed to Microgravity" (US-3). Hardware was developed for all three experiments to fit into the European Space Agency's "Type I/O" or "Type II/O" containers.
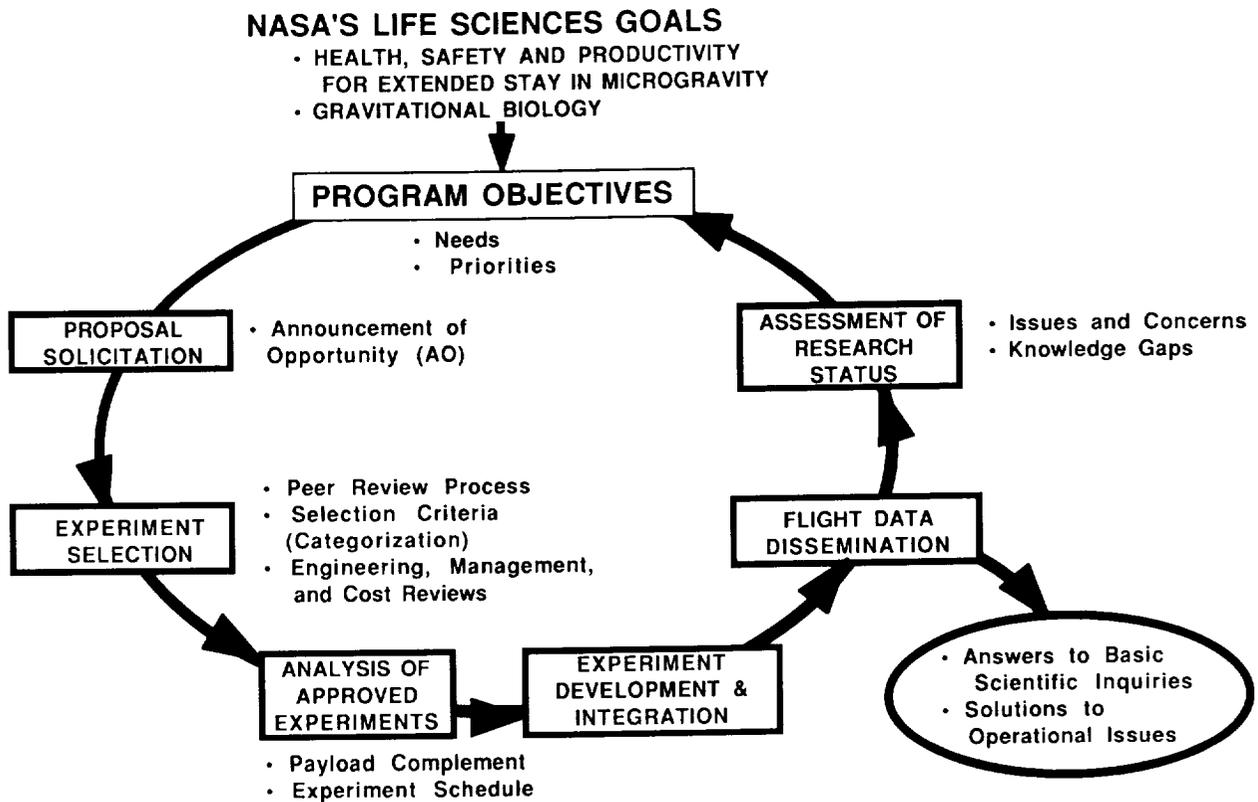
Figure 1. Space Life Sciences research planning process.
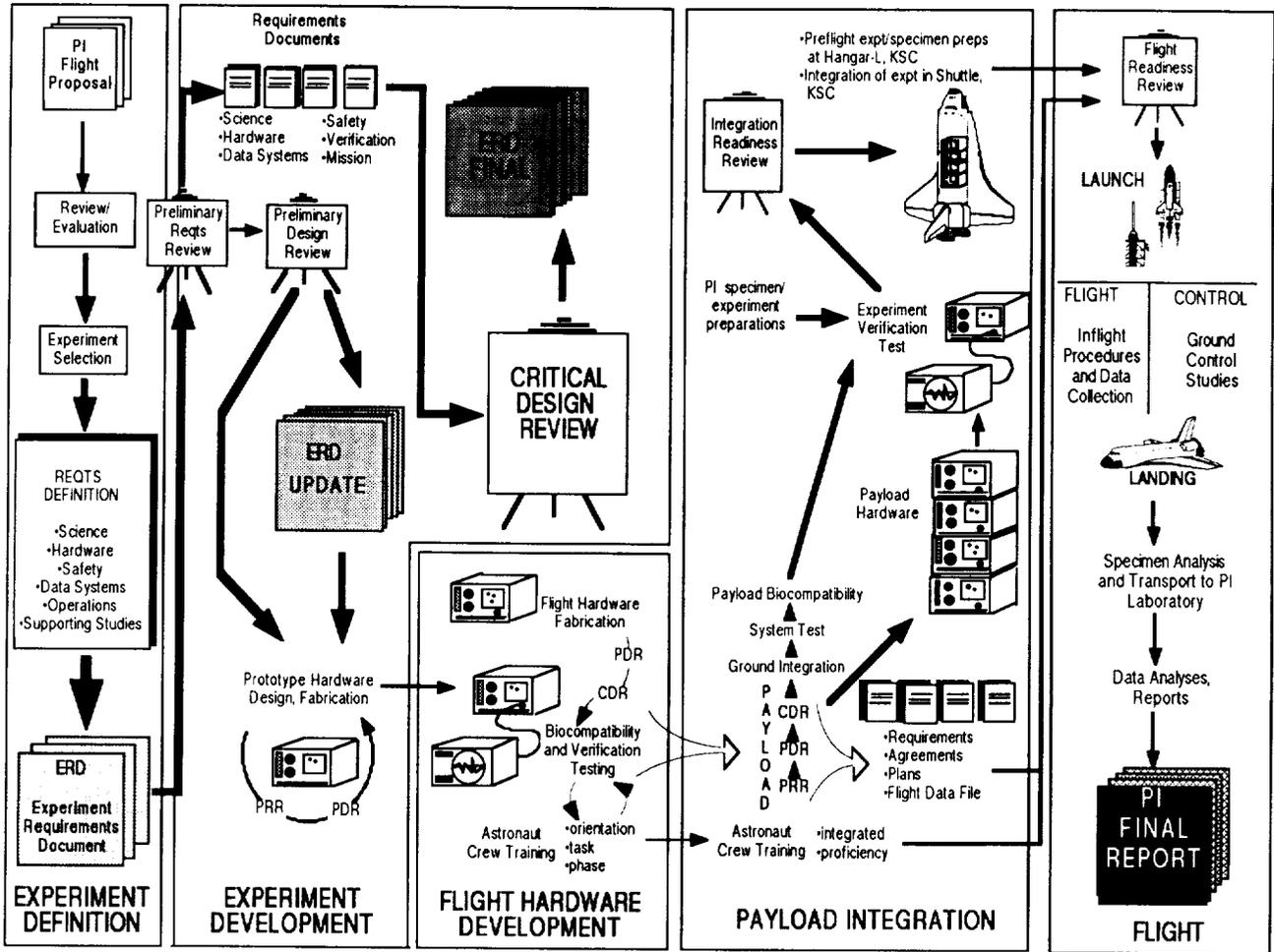
Figure 2. Phases of an experiment, from Proposal to Flight.

MILESTONES AND EXPERIMENT DEVELOPMENT SHOWN AS MONTHS BEFORE LAUNCH
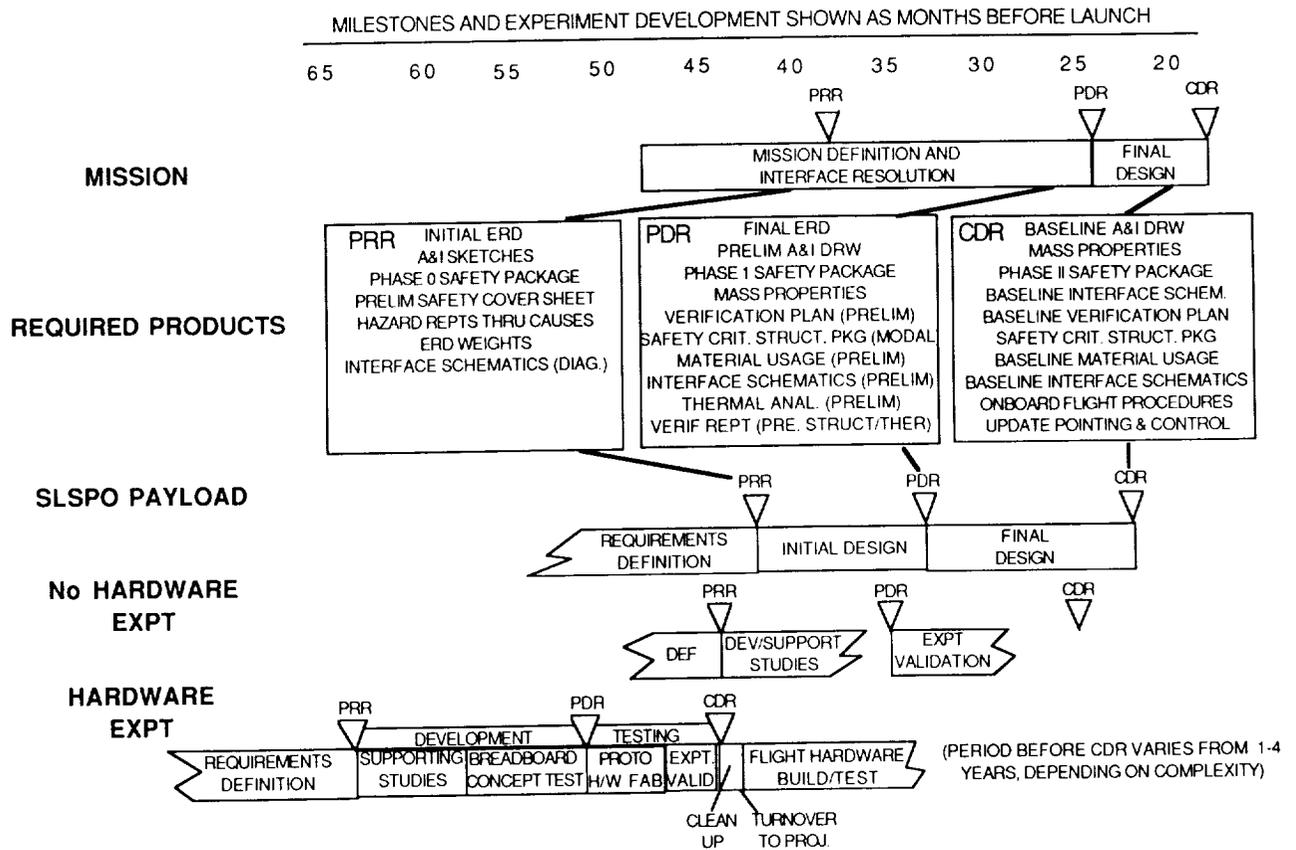
Figure 3. Timelines for an experiment's development.

SECTION IV

ENVIRONMENTAL INTERACTIONS

I

Session E1: PLASMA INTERACTIONS - I: PLANNED SPACE EXPERIMENTS

Session Chair: N. T. Grier

THE SOLAR ARRAY MODULE PLASMA INTERACTIONS EXPERIMENT (SAMPIE):
Science and Technology Objectives

G. Barry Hillard

NASA Lewis Research Center
Cleveland, Ohio 44135

ABSTRACT

The Solar Array Module Plasma Interactions Experiment (SAMPIE) is an approved NASA shuttle space flight experiment to be launched in July 1993. The SAMPIE experiment is designed to investigate the interaction of high voltage space power systems with ionospheric plasma. To study the behavior of solar cells, a number of cell coupons, representing technologies of current interest, will be biased to high voltages to characterize both negative potential arcing and positive potential current collection. Additionally, various theories of arc suppression will be tested by including several specially modified cell coupons. Finally, SAMPIE will include experiments to study the basic nature of these interactions. This paper describes the rationale for a space flight experiment, the measurements to be made, the significance of the expected results, and the current design status of the flight hardware.

BACKGROUND

Traditionally, space power systems in Low Earth Orbit (LEO) have operated at low voltages and have not suffered from the effects of plasma interactions. High power systems now under development for space applications will operate at high end-to-end voltages in order to minimize array current. The emergence of such systems is motivated primarily by a desire to save weight. Since the resistance of the necessary cabling is a strongly decreasing function of mass per unit length and cable losses are proportional to current squared, it is desirable to operate at high voltages and low currents. A further consideration is the reduced effect of magnetic interactions (torque and drag) that will follow from low current operation.

While high voltage systems are clearly desirable to the power system designer, they suffer the drawback of interacting with the ionospheric plasma (1,2) in two different ways. First, conducting surfaces whose electrical potential is highly negative with respect to the plasma undergo breakdown and arcing. Such arcing not only damages the material but results in current disruptions, significant electromagnetic interference (EMI), and large discontinuous changes in the array potential. For arrays using traditional silver-coated interconnects, a threshold potential for arcing of about -230 volts relative to the plasma is believed (3) to exist. There are theoretical arguments (4) supported by limited ground test results (5) that different metals will arc at different thresholds. Since new solar cell designs are emerging using copper traces, it is important to determine arcing thresholds, arc rates, and arc strengths for a variety of materials exposed to space plasma.

For solar arrays or other surfaces which are biased positive with respect to the plasma, a second effect occurs. Such surfaces collect electron current from the plasma resulting in a parasitic loss to the power system. Since the mass of electrons is much less than ions, the magnitude of current collection is much greater for surfaces with positive bias. At bias potentials greater than about 200 volts, sheath formation causes the entire surrounding surface, normally an insulator, to behave as if it were a conductor. This effect, called "snapover," results in large current collection even from a very small exposed area. Besides producing a power loss, this current will significantly affect the potentials at which different parts of the array will "float." Depending on the way the power system is grounded, this in turn will affect the equilibrium potentials of various spacecraft surfaces with respect to the plasma.

Two previous flight experiments involving standard silicon arrays, PIX I and PIX II (1,2) have shown many differences between ground tests and behavior in space. For arcing, arc rates in space were quite different and generally higher than in ground tests. For parasitic current collection, the current versus bias voltage curves obtained in space not only differed radically from the ground tests but differed depending on whether the data was taken with the array exposed to spacecraft ram or wake. It is necessary, therefore, that the behavior of various solar cell technologies be established with a suitable in-space test.

In this paper, we have only briefly reviewed the background and justification for SAMPIE since this has been presented previously (6). We will present the status of the design and a discussion of the selected experiments to be done.

OBJECTIVES

The general objective of SAMPIE is to investigate, with a Shuttle-based space flight experiment and relevant ground-based testing, the arcing and current collection behavior of materials and geometries likely to be exposed to LEO plasma on high voltage space power systems. There are seven specific objectives of the SAMPIE experiment:

1. For a selected number of solar cell technologies, determine the arcing threshold and arc rates and strengths. At a minimum, the solar cells selected for flight must include:

    a. A sample array made of traditional silicon solar cells. This will provide a baseline for comparison with past experiments.

    b. A sample array using APSA, the Advanced Photovoltaic Solar Array.

    c. A sample array using current space station solar cell technology.

2. For these sample arrays, determine the plasma current collection characteristics.

3. Propose, demonstrate in ground tests, and fly an arc mitigation strategy; i.e., modifications to standard

interconnect design which may significantly improve the arcing threshold.

4. Design simple metal/insulator mockups to allow the dependance of current collection on exposed area to be studied with all other relevant parameters controlled.

5. Design a simple arcing experiment to test the dependance of arcing threshold, arc rates, and arc strengths on the choice of metal with all other relevant parameters controlled.

6. Design, test, and fly simple controlled experiments to study basic phenomena related to arcing and its effects. Added on a space-available basis subject to time and resource constraints, these may include such things as:

    a. Arcing from anodized aluminum using alloys and anodization processes typical of ones being considered for use on large space structures.

    b. Arcing from pinholes in Indium-Tin oxide (ITO) coated conductors or from biased conductors covered with strips of ITO.

    c. Sputtering and degradation of metals or metal covered insulators biased to high negative potential in the atomic oxygen environment of LEO.

7. Measure a basic set of plasma parameters to permit data reduction and analysis. An additional requirement to aid data reduction is to provide timely flight data (such as the Shuttle orientation, and times of thruster firings) relevant to SAMPIE flight conditions.

APPROACH

SAMPIE will consist of a metal box with an experiment plate fixed to the top surface. It will mount directly to the top of a Hitchhiker-M carrier. A power supply will bias the solar cell samples and other experiments to DC voltages as high as +700 volts and -700 volts with respect to shuttle ground. When biased negative, suitable instruments will detect the occurrence of arcing and measure the arc-rate as a function of bias voltage. For both polarities of applied bias, measurements will be made of parasitic current collection versus voltage. Other instruments will measure the

degree of solar insolation, plasma electron density and temperature, and monitor the potential of the shuttle with respect to the plasma. Shuttle operational logs will be relied upon for detailed information about the orientation of the experiment with respect to the vehicle's velocity vector as well as times and conditions of thruster firings.

In a simplified description of the experiment, one sample is biased to a particular voltage for a preset time while measuring arcing and current collection data. A set of plasma diagnostics is then taken and the procedure repeated at the other bias voltages until all measurements are completed. Vehicle orientation is critical since ram and wake effects are known to be significant. SAMPIE will request control of the orbiter orientation such that one entire set of measurements is made with the payload bay held in the ram direction and a second set with the bay in the wake.

DESIGN STATUS

Since SAMPIE was originally designed to be deployed on a 15 meter collapsible tube mast of ESA
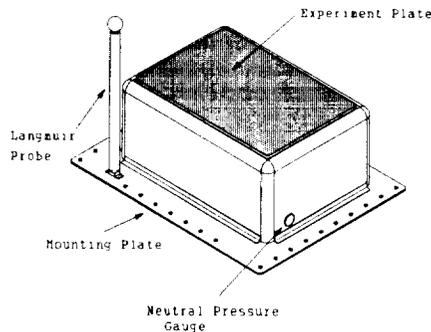
Fig. 1 External view of SAMPIE package

design (6), it has been severely constrained in mass. As a result, although the current baseline is for direct mounting to the Hitchhiker carrier, the package remains quite compact. Figures 1 and 2 show the basic package.

Figure 3 shows the proposed layout of the experiment plate. To meet objectives 1 and 2, which require extensive solar cell testing, a number of cell coupons are provided.
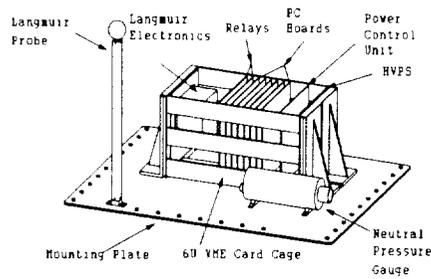
Fig. 2 Internal view of SAMPIE package

a.   A baseline for comparison is provided by including a small 9-cell coupon of standard technology silicon 2 cm by 2 cm cells. This is the technology that has been used exclusively in the U.S. space program to date. It was flown on PIX I and PIX II as well as being the subject of extensive ground based testing and will provide a basis for continuity with past results. A second coupon of standard cells is shown surrounded by a metal guardring. this is simply a metal structure which can be biased independently of the cell coupon and is designed to test the effect of a large surrounding solar array. NASCAP/LEO will be used to determine the appropriate bias voltages for each bias applied to the coupon.
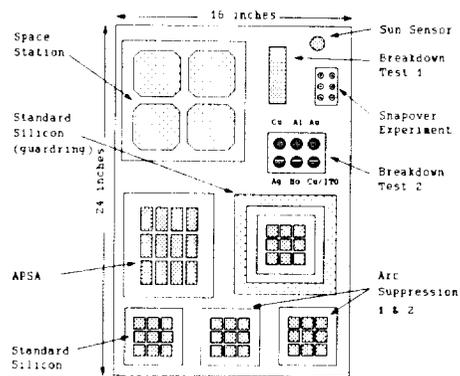
Fig. 3 SAMPIE experiment plate

b. A 4-cell coupon of 8 cm by 8 cm space station cells, having copper interconnects in the back will allow a test of this technology.

c. A 12-cell coupon of 2 cm by 4 cm APSA cells will test the behavior of this relatively new, very thin (60 micron) technology. APSA is normally a flexible blanket mounted in an external frame while the original intent on SAMPIE was to mount the cells directly to the stiff experimental plate. While it is highly likely that the plasma interactions encountered by the cell array will be the same, this point has yet to be proven. Two coupons, one mounted rigidly and the other in the flexible, baseline design, will be subjected to extensive ground testing. If there are clearly no differences, it will be easier to fly the rigidly mounted assembly. If differences are found, every attempt will be made to design a flight-qualifiable mounting scheme for the flexible array segment.

There are several experiments designed to test basic theories of arcing and current collection.

a. The first of the two breakdown tests shown in figure 3 will explore the hypothesis that negative potential arcing is a special case of the classical vacuum arc (7). With geometry and test conditions controlled, only the composition of the metal will be varied. The resulting family of arc rate versus bias voltage curves will give considerable insight into the basic nature of the arcing process. The particular choice of metals is based on current and anticipated importance to space technology.

b. The second breakdown test consists of a single sample of anodized aluminum. There is considerable concern that this material undergoes dielectric breakdown and arcing when biased to high voltages (8). The particular alloy and anodization process are chosen to be identical with structural material currently baselined for Space Station Freedom.

c. To study current collection and snapover, we include six 1-cm diameter copper disks covered with 5 mil kapton. Each has a pinhole in the center with hole sizes tentatively chosen as .1 mm, .3 mm, .5 mm, .7 mm, 1 mm, and 1.5 mm. The resulting family of current versus applied bias curves will be compared with pre-

dictions of NASCAP/LEO and other theoretical treatments.

Several arc suppression techniques are under investigation as part of our ground based testing. These generally follow from the work of Katz et. al. (9) on the SPEAR program which showed that inbound ions striking the junction of insulator, metal, and plasma, sometimes called the triple point, result in secondary emission and arcing.

a. The first technique we will test follows from recent NASCAP/LEO (10) modeling done in support of Space Station Freedom. The results indicate that simply extending the cover slides to cover a larger portion of the gap between cells is sufficient to choke off most of the ion current. We will test this idea with specially modified 2 cm by 2 cm silicon cells since space limitations would make a second space station coupon difficult to accommodate.

b. The second technique we will test was developed by Physical Sciences Inc. under a still open SBIR contract (11). This work has shown that a major factor in arcing is ion bombardment of excess adhesive which is inevitably present in the gap between cells. Current plans are for a coupon of silicon 2 cm by 2 cm cells to be sent to PSI, subjected to their newly developed cleaning process, and returned for incorporation into SAMPIE.

SUMMARY

The SAMPIE flight experiment is the first orbited space power system - plasma interaction experiment since PIX II and is by far the most ambitious to date. Besides testing two emerging solar cell technologies, it will explore the viability of several arc suppression techniques. Using controlled experiments, it will provide basic data on arcing and current collection which can be used to validate and extend existing models and theories. SAMPIE will be designed and built in a highly modular way that will have easy reflight capability in mind. To this end, it can serve as a test-bed for future technologies.

REFERENCES

1. Grier, N.T. 1983, "Plasma Interaction Experiment II (PIX II): Laboratory and Flight Results", SPACECRAFT ENVIRONMENTAL INTERACTIONS TECHNOLOGY 1983, NASA CP-2359, pp. 333-347

2. Grier, N.T. and Stevens, N.J. 1978, "Plasma Interaction Experiment (PIX) Flight Results", SPACECRAFT CHARGING TECHNOLOGY 1978, NASA CP-2071, pp. 295-314

3. Ferguson, D.C. 1986, "The Voltage Threshold for Arcing for Solar Cells in LEO - Flight and Ground Test Results", NASA TM-87259.

4. Jongeward, G.A. et. al. 1985, "The Role of Unneutralized Surface Ions in Negative Potential Arcing", IEEE TRANS. NUCL. SCI., vol. NS-32, no. 6, Dec., pp 4087-4091

5. Snyder, D.B. 1986, Private Communication

6. Ferguson, D.C. "SAMPIE - A Shuttle-Based Solar Array Arcing Experiment", PROCEEDINGS OF THE SPACECRAFT CHARGING TECHNOLOGY CONFERENCE, Monterey CA, 31 October - 3 November 1989

7. Hillard, G.B., "Negative Potential Arcing: Current and Planned Research at LeRC", PROCEEDINGS OF THE SPACECRAFT CHARGING TECHNOLOGY CONFERENCE, Monterey CA, 31 October - 3 November 1989

8. Carruth, R. 1990, Private Communication

9. Katz and Cooper, U.S. patent 4835841

10. Ferguson, D.C. and Chock R.R., "Floating Potentials of Space Station Freedom with Present and Modified Solar Cell Designs: Analysis of Current Collection by SSF Solar Cell", to be published

11. Upschulte, B.L. et. al., "Significant Reduction in Arc Frequency of Negatively Biased Solar Cells: Observations, Diagnostics, and Mitigation Techniques", to be published in THE PROCEEDINGS OF THE "ELEVENTH SPACE PHOTOVOLTAIC RESEARCH CONFERENCE (SPRAT XI)", Cleveland, OH, May 7 - 9 1991

NASCAP/LEO SIMULATIONS OF SHUTTLE ORBITER CHARGING
DURING THE SAMPIE EXPERIMENT

Ricaurte Chock
NASA Lewis Research Center
MS 301-3
21000 Brookpark Rd., Cleveland, OH 44135

ABSTRACT

The electrostatic charging of the Shuttle Orbiter during the operation of the Solar Array Module Plasma Interaction Experiment (SAMPIE) has been modeled using the NASCAP/LEO computer code. The SAMPIE experiment, scheduled to be flown in the shuttle payload bay in 1993, consists of an array of various solar cells representing the present technologies. The objectives of the experiment are to investigate the arcing and current collection characteristics of these cells when biased to high potentials in a low Earth orbit (LEO) plasma. NASCAP/LEO (NASA Charging Analyzer Program/Low Earth Orbit) is a 3-D code designed to simulate the electrostatic charging of a spacecraft exposed to a plasma at low earth orbit or ground test conditions. At its most extreme configuration, with the largest array segment of the SAMPIE experiment biased +600 V with respect to the Orbiter and facing the ram direction, the computer simulations predict that the Orbiter's potential will be approximately -20 V with respect to the plasma.

I. INTRODUCTION

NASCAP/LEO simulations comparing ground test results with low earth orbit conditions have highlighted the difficulties encountered when trying to extrapolate solar array behavior under LEO conditions from vacuum chamber experiments. NASCAP/LEO (NASA Charging Analyzer Program/Low Earth Orbit) is a 3-D code designed to simulate the electrostatic charging of a spacecraft exposed to a plasma at low earth orbit or ground test conditions (ref. 1). Using this code it has been found that small changes in cell geometry, such as allowing a cell cover glass overhang of 6 mils, will greatly impact the cell's current collection behavior (ref. 2). In order to better understand such

behavior, actual flight experiments are needed. One of these is the Solar Array Module Plasma Interaction Experiment (SAMPIE), scheduled to be flown in the Orbiter bay in 1993 (ref. 3).

The SAMPIE experiment consists of an array of various solar cells, representing the present technologies. The objectives of the experiment are to investigate the arcing and current collection characteristics of these cells when biased to high potentials in a LEO plasma. These collection and arcing measurements will be made with the cells biased up to ± 600 V facing the ram and wake directions.

In LEO the Orbiter's potential will change so that the net current to the Orbiter from the plasma is zero. The potential at which this occurs is defined as the Orbiter's floating potential. With the SAMPIE cells biased to +600 V and facing the ram direction there is a possibility that the Orbiter's floating potential will be driven highly negative to balance and cancel out the incoming electron current.

In order to better design the experiment so as to avoid possible arcing damage to the Orbiter, NASCAP/LEO was used to model the Orbiter's electrostatic charging and obtain possible Orbiter floating potentials under different experimental configurations. From the available data (ref. 4), we can infer a floating potential of about -70 V for Skylab. This floating potential did not cause any problems during Skylab operations so -70 V was used as an acceptable floating potential.

II. NASCAP/LEO SIMULATION

First a finite element model of the Orbiter was created (see Fig. 1) using PATRAN® (a registered trademark of PDA Engineering), a commercially

available 3-D mechanical computer-aided engineering software system. The Orbiter is modeled as being a dielectric object whose only conductors are its main engine and thruster nozzles. The SAMPIE experiment is placed in the approximate center of the bay.

The SAMPIE experiment itself is modeled as a box with the dimensions (.45x.45x.25 m) of the actual experiment and all of the top plate defined as a conductor. The top plate has no individual features such as solar cell assemblies or other experiments. This is because NASCAP/LEO's resolution can't distinguish individual features on the plate and still include the Orbiter in its computational grid.

NASCAP/LEO is a modular code. Each module is a program, or collection of programs, which solve a particular aspect of the spacecraft charging problem. A call to the CURRENT module, for example, will calculate currents from the plasma to the spacecraft. Other available modules are RDOPT and IPS. In the RDOPT (Read Options) module the user can input parameters such as plasma temperature and density, spacecraft speed and conductor potentials among others. The IPS (Initial Potential Specification) module calculates the electrostatic potentials of the spacecraft's surfaces and it's surrounding space environment. To run any given simulation one calls each of the modules individually.

The floating potential of the Orbiter with the SAMPIE experiment in operation was calculated using the RDOPT, IPS, and CURRENT modules. The procedure is straightforward. One performs several NASCAP/LEO code runs varying the Orbiter's potential until the net current to the spacecraft is negligible. The potential at which this occurs is then taken as the Orbiter's floating potential.

This procedure would only take into account sheath generated particles but by including three QUICK, CHARGE, POTENT cycles in the simulation we can take into account ambient particles as well. QUICK, CHARGE, and POTENT are other modules available from the NASCAP/LEO code. All the computer runs for this paper were done on a Celerity 1200 mini computer running Accel 4.2 UNIX. Further details on this procedure or about the NASCAP/LEO modules can be found in the NASCAP/LEO User's Guide (ref. 5). A sample of NASCAP/LEO input is shown in Fig. 2.

At the end of each simulation a CURRENTS utility, not to be confused with the CURRENT module, is run. The output from CURRENTS consists of the electric current values to the Orbiter/SAMPIE surfaces. This output is divided into material and conductor surfaces. In the present simulations one can read individual current values to the Orbiter, thruster nozzles, bay area, body (wings, empennage, cabin), the top of the bay doors, and the SAMPIE plate as well as the total current to the Orbiter/SAMPIE object.

In this paper, the simulations consist of the four experimental configurations listed below:
Case 1: SAMPIE in the Orbiter's ram, biased to +600 V.
Case 2: SAMPIE in the Orbiter's ram, biased to -600 V.
Case 3: SAMPIE in the Orbiter's wake, biased to +600 V.
Case 4: SAMPIE in the Orbiter's wake, biased to -600 V.
All SAMPIE biases are with respect to the Orbiter potential. Orbiter potentials are with respect to plasma ground.

III. FLOATING POTENTIAL DETERMINATION AND SIMULATION RESULTS

Case 1 is the most critical. With the SAMPIE experiment biased highly positive and facing the incoming ram particle flux, one can expect SAMPIE to draw large negative currents from the plasma. To cancel out this current the Orbiter will charge negatively in order to repel electrons and attract ions from the plasma. Depending on the magnitude of these currents the Orbiter may charge highly negative, thus exceeding safety limitations and interfering with the successful completion of the experiment.

When SAMPIE is biased negative and facing the ram direction it will collect ions proportional to the ram ion flux on its surface. Ram ion flux for LEO is in the order of 10^{-5}-10^{-3} A/m^2 (ref. 5) so one would expect small currents for case 2, thus a low floating potential.

One would also expect low floating potentials for cases 3 and 4 because of the reduced plasma density due to wake effects. A spacecraft flying through the plasma at a typical LEO velocity of 7500 m/s creates a region behind it in which both electron and ion densities are reduced in comparison with an undisturbed plasma. This spacecraft velocity is about six times larger than the ion thermal velocity (using a .1 eV oxygen ion) so a spacecraft would travel a distance equivalent to several of its own radii before the ions could fill in the region behind it. Electrons are more mobile and can fill in the
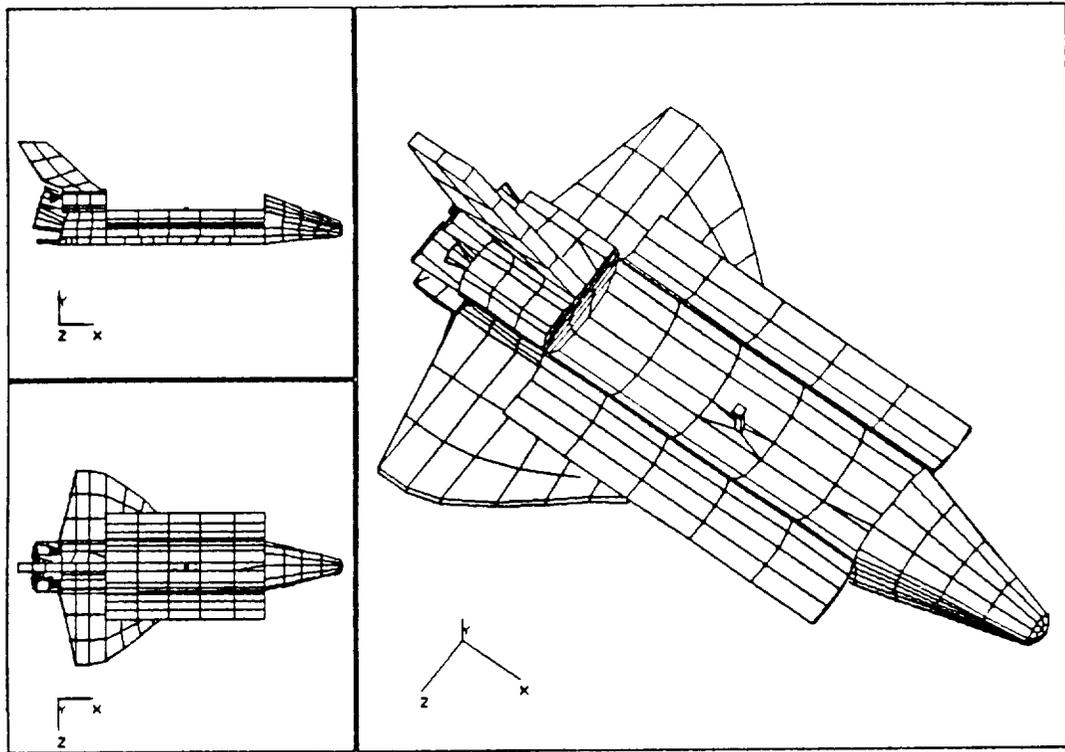
Fig. 1: NASCAP/LEO model of the Shuttle orbiter with the SAMPIE experiment on the center of the bay.

```
rdopt 5                         # Read computational options
    temperature .1              # Plasma temperature (eV)
    rbo 1.1e11                  # Plasma density (#/m³)
    errlim 0.1                  # convergence error parameter
                                  for IPS module
    pcond 1 -200                # Sets conductor 1 to -200 V
    bias 2 600                  # Biases conductor 2 600 V positive with respect
                                  to conductor 1
    satvel 7500 0 0             # Object speed (x y z) m/s
    ionmass 16 amu             # Oxygen ion mass
    sheath boundary 1          # Sheath defined at the 1 V contour
    end                        # End of computational options
wake                           # Calculates reduced ion densities due to wake
                                  effects
ips 5                          # Initial Potential Specification
    all matl -.3               # Assigns a surface potential of -.3 V to material
                                  matl as an initial guess
    end                        # End of IPS options
quick
charge
potent
quick
charge                         # 3 QUICK, CHARGE, POTENT cycles
potent
quick
charge
potent
current                        # Calculates currents to spacecraft
end                            # End of NASCAP/LEO run
```

Fig. 2: NASCAP/LEO sample input.

657

wake region more rapidly. However this is limited
by the space charge of electrons already present in
the wake, so for most of the wake region the
electron and ion densities are comparable (ref. 5).
Results of measurements done by Murphy et al.
(ref. 6) from within the Orbiter bay indicate a
decrease of 3 orders of magnitude in electron
density as a conservative estimate for the near wake
region.

Case 1: SAMPIE in the ram, biased +600 V with
respect to the Orbiter

As the Orbiter potential increases from 0 V to
-400 V we see that currents to the Orbiter body
and the top of the bay doors are negligible
compared to the other currents, see Fig. 3. The

Fig. 4: Case 1, currents to spacecraft vs. Orbiter
potential.
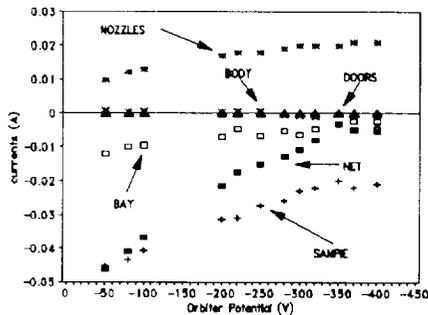
Fig. 3: Case 1, currents to spacecraft vs. orbiter
potential. NET is the net current to Orbiter.

door tops are shielded by SAMPIE's sheaths most
of the time. The Orbiter/SAMPIE object collects
current mainly through the Orbiter nozzles, the
SAMPIE plate, and the bay area. The majority of
the ions are collected by the nozzles while the
electrons are collected by the experimental plate and
the bay. The bay area is a dielectric surface but the
fact that it is moving into the ram and that
SAMPIE's sheaths focus charge into the bay will
allow it to collect charge up to approximately
-2.4 mA, from then on it will not collect larger
currents.

At an Orbiter ground potential of -350 V the
currents collected by the nozzles cancel out the
current collected by SAMPIE, see Fig. 4. However
at this potential the bay is charged up to about
-3.4 mA. There is no positive charge large enough
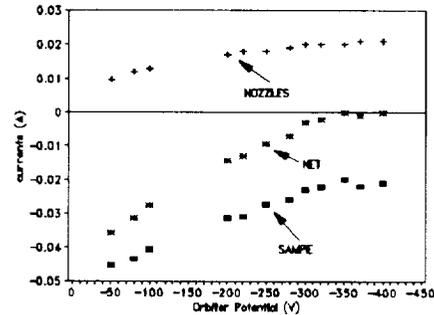
to cancel the bay charge. The positive current
collected on the body is on the 10^{-4}A range and
these currents will not flow through the dielectric
body to the bay. From these results one might infer
a floating potential of about -350 V for case 1.

The floating potential may not be as large in reality
because only a small area of the plate (some cell
interconnects) will be biased to +600 V relative to
the Orbiter instead of the whole plate surface as the
simulation assumes. These cell interconnects would
then be the effective collecting area. Assuming that
the sheath through which SAMPIE current is
collected scales proportionally to the conducting
area of the plate we can scale the currents to
SAMPIE by reducing the plate area. One can use
these new currents to obtain a better estimate of the
Orbiter/SAMPIE floating potential by reducing the
currents to SAMPIE by a factor of actual collecting
area vs. total experimental plate area. This should
provide an estimate of the actual current collected
by the experiment. These currents are then plotted
to obtain a new I/V curve from which one can
deduce a floating potential.

First it is important to verify if reducing the area
of the SAMPIE plate on the simulation by a given
factor reduces its current collection by a similar
factor. Upon inspection of the Orbiter/SAMPIE
object it can be seen that one can reduce the area of
the experimental plate to half of its original value
and still be within the margins of NASCAP/LEO's
grid resolution. One can use this second
Orbiter/SAMPIE model in a simulation and
compare its currents to the original model. If the

658

currents to this new SAMPIE object are approximately half their original value the approach is correct.

The new Orbiter/SAMPIE object will be referred to as SAMPIE2, where SAMPIE2's experimental plate is one half the area of SAMPIE1's plate model. In the SAMPIE2 simulation at high voltages the ratio of currents SAMPIE2/SAMPIE1 is very near to .5 as was previously assumed while the currents to the nozzles remains the same (see Fig. 5). At low

Fig. 6: Case 1 floating potential determination.

Fig. 5: Case 1, ratio of currents to spacecraft. SAMPIE2 vs. SAMPIE1 simulations.

voltages the ratio is between .4 and .5 which is not in bad agreement with our assumption. It can also be seen that for an Orbiter ground potential smaller than -300 V the ratio of current to the bay is .1. So if one reduces the collecting area, the bay currents will become negligible. One can thus be reasonably confident in scaling the currents to the SAMPIE plate by an appropriate area factor.

Assuming a worst case in which the whole surface area of the cell will act as a conductor, i.e. the cells will be "snapped over", a likely possibility with this high positive bias. With four Space Station Freedom type solar cells as the base line there is a total surface area of about $2.48 \times 10^{-2} m^2$. The ratio of this area to the original SAMPIE plate area is approximately .124. Reducing the SAMPIE currents by .124 and graphing them (see Fig. 6) one finds a floating potential of about -20 V.

The NASCAP/LEO simulations therefore predict that when the SAMPIE experiment is biased +600 V with respect to the Orbiter and facing the ram, the shuttle's floating potential will be in the range of -20 V. One also sees that for this type of experiments there is a limiting size to the

experimental array beyond which using the shuttle becomes impractical and possibly hazardous. The simulation shows that a plate .2 m^2 in area can not be biased to high positive voltages without driving the shuttle's floating potential highly negative.

Case 2: SAMPIE in the ram, biased -600 V with respect to Orbiter

One proceeds in the same manner as described above. Since this case is not expected to be critical the original Orbiter/SAMPIE model (SAMPIE1) may be used.

One expects low currents for this configuration and the simulations bear this out. The Orbiter ground voltage was changed from -50 V to +300 V. At all voltages SAMPIE current collection was small, around 6.5×10^{-4} A to 8.2×10^{-4} A. The sheath is localized around the SAMPIE box and does not affect current collection as in case 1. In this case the Orbiter connects to the plasma through the nozzles as if no experiment were present therefore the net current to the Orbiter consists of the nozzle's current.

The floating potential for case 2 can be calculated to be in the range of -10 V, see Fig. 7.

659
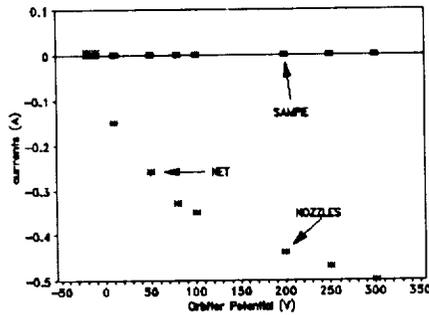
Fig. 7: Case 2 floating potential determination.

CASE 3: SAMPIE in the wake, biased +600 V with respect to Orbiter

SAMPIE collects electron current but only on the order of 10^{-5}A which is small when compared to the 10^{-3}A to 10^{-2}A nozzle current. So once again the net current to the Orbiter is the current to the nozzles. This current is zero for an Orbiter potential between -5 V and -2 V (see Fig. 8).
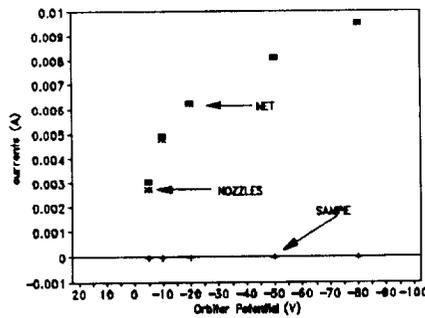
Fig. 8: Case 3 floating potential determination.

CASE 4: SAMPIE in the wake, biased -600 V with respect to Orbiter

As before the only connection to the plasma is through the Orbiter nozzles. The SAMPIE plate in the wake is a poor ion collector. In Fig. 9 it may be seen that the Orbiter will float at about 0 V. The CURRENTS output seems to indicate it will float slightly positive between 0 V and +1 V.
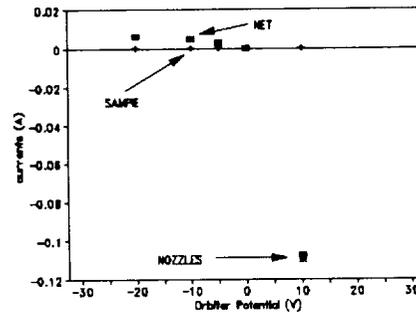
Fig. 9: Case 4 floating potential determination.

III. CONCLUSIONS

The NASCAP/LEO simulations predict that while the operation of the SAMPIE experiment will have an impact on the Orbiter's floating potential, it will not be a serious one. A worst case of -20 V has been predicted which is within the -70 V mentioned before as an acceptable floating potential. They also indicate possible limitations in similar experiments, for example, the same experiment with an active array collecting area of .2 m^2, biased highly positive, would drive the shuttle's potential to unacceptably large negative voltages. The SAMPIE experiment will not have this problem because the biased area is small and the platform upon which it is mounted, i.e. the Orbiter, has good contact with the plasma via the thruster nozzles. However, it is imperative that the Orbiter nozzles not be in the Orbiter's wake during the SAMPIE experiment for this will decrease the nozzles electrical contact with the plasma. Other high voltage experiments mounted on platforms which do not have a large exposed conductive area may charge up to large potentials which may then interfere with the experiment's operation.

REFERENCES

1. Mandell, M.J., Katz, I., "High Voltage Plasma Interaction Calculations Using NASCAP/LEO", AIAA-90-0725, 28th Aerospace Sciences Meeting, Reno, Nevada, Jan 8-11 1990.

2. Chock, R., "NASCAP/LEO Simulations of Space Station Cell's Current Collection", unpublished.

3. Hillard, B.G., "The Solar Array Module Plasma Interaction Experiment (SAMPIE)", SOAR '90 Proceedings NASA Conference Publication 3103 Vol. II, Albuquerque, New Mexico, Jun 26-28 1990.

4. Woosley, A.P., Smith, O.B., Nassen, H.S., "Skylab Technology - Power Systems", THE SKYLAB RESULTS, AAS Publication Office, Tarzana, California, 1975, 559.

5. Mandell, M.J., Davis, V.A., "User's Guide to NASCAP/LEO", Draft

6. Murphy, G., Pickett, J., D'Angelo, N., Kurth, W.S., "Measurements of Plasma Parameters in the Vicinity of the Space Shuttle", Planet. Space Sci., 34, 10, 993

PASP PLUS: AN EXPERIMENT TO MEASURE SPACE-ENVIRONMENT EFFECTS ON PHOTOVOLTAIC POWER SUBSYSTEMS

Donald A. Guidice
Phillips Laboratory, Hanscom AFB, MA 01731

Abstract: The Photovoltaic Array Space Power Plus Diagnostics experiment (PASP Plus, for short) has been accepted as part of the APEX mission payload aboard a Pegastar satellite to be orbited by a Pegasus launch vehicle in late 1992. The mission's elliptical orbit (190 nmi perigee, 1000 nmi apogee, nominal) will allow us to investigate both space plasma and space radiation effects. PASP Plus will have eleven types of solar arrays and a full complement of environmental and interactions diagnostic sensors. Measurements of space-plasma interactions on the various solar arrays will be made at large negative voltages (to investigate arcing parameters) and at large positive voltages (to investigate leakage currents) by biasing the arrays to various levels up to -500 and +500 Volts. The long-term deterioration in solar array performance caused by exposure to space radiation will also be investigated; radiation dosage will be measured by an electron/proton dosimeter included in the environmental sensor complement. Experimental results from PASP Plus will help establish cause-and-effect relationships and lead to improved design guidelines and test standards for new-technology solar arrays.

INTRODUCTION

Future space system operations will require higher powered photovoltaic subsystems to carry out more ambitious missions. Standard planar-array power sources using small *silicon* solar cells in low-voltage configurations (not much above 28 Volts) could be superseded by larger cells (to reduce array costs), different materials (with higher efficiencies), and higher voltage levels (to minimize cable losses or reduce cable weight). However, before any space-system designers will commit to using new technologies or configurations, systematic investigations of the effects of the space environment on the performance of advanced solar arrays must be made. The Photovoltaic Array Space Power Plus Diagnostics (PASP Plus) experiment provides a means for carrying out the needed investigations.

Originally, as described in a SOAR '89 paper (Ref. 1), the objectives of the PASP Plus experiment were limited to the investigation of the effects of space-plasma interactions on high-voltage solar array operations at low altitudes. However, in early 1990 the Space Test Program of Space Systems Division (SSD) offered the PASP Plus experiment a flight on a Pegastar satellite boosted into orbit by a Pegasus launch vehicle. PASP Plus became part of the APEX (Advanced Photovoltaic and Electronics Experiments) mission set up to fly PASP Plus and two small "radiation effects on electronics" experiments, CRUX/CREDO and FERRO. A Spaceflight Plan for APEX was approved by Hq USAF on 3 October 1990. Because of the enhanced opportunity provided by APEX—an elliptical [350 km by 1850 km] near-polar [i = 70°] orbit with a one to three year lifetime, it was decided to broaden the scope of PASP Plus to include the investigation of the effects of space radiation dose on long-term solar array performance. We also added diagnostic instruments appropriate to PASP Plus's new scope and mission profile. This paper is intended to describe the features of the new PASP Plus experiment.

PASP PLUS OBJECTIVES

The objectives of the expanded-scope PASP Plus experiment are:

   a. To measure the plasma "leakage" current for many different arrays subjected to positive biasing levels up to +500 V, simulating array operation at high positive voltages.

   b. To measure the arcing parameters for many different arrays subjected to negative biasing levels up to -500 V, simulating array operation at high negative voltages.

c. To measure the long-term deterioration in the electrical performance of many different types of solar arrays exposed to space radiation.

d. To provide an opportunity to test various new photovoltaic technologies (new materials and/or designs) in the real space environment.

e. To establish cause-and-effect relationships between array performance and environmental conditions.

PASP PLUS INSTRUMENTATION

The experiment consists of four kinds of equipment:

a. a set of test arrays, several of which are divided into biased and unbiased segments.

b. experiment-control instrumentation capable of creating array or spacecraft conditions under which measurable environmental interactions will occur.

c. interactions-measuring instrumentation that will quantify what happens when the ambient or created conditions impact the performance of the test arrays.

d. diagnostic sensors to measure the ambient space-environment conditions affecting array performance.

Solar Array Complement   To maximize the applicability of the experiment, a variety of conventional and advanced-concept solar arrays were included in the PASP Plus array complement. As shown in Table I, eleven different solar arrays will be investigated on PASP Plus. Other than array #1 (which serves as a standard), the decisions on array selection made for PASP Plus were based on the array's potential utilization on future DoD or NASA spacecraft missions.

TABLE I. PASP PLUS SOLAR ARRAYS

ARRAY	CELL TYPE	DESCRIPTION	SIZE(in x in)	BIASED SEGMENTS
1	Si	2cm x 2cm, BSF	10 x 20	2 of 3
2	Si	8cm x 8cm, 8-mil WTC (for Space Station)	8 x 9.5	1 of 1
3	GaAs/Ge	4cm x 4cm, 3.5-mil	10 x 20	2 of 3
4	GaAs/Ge	4cm x 4cm, 7-mil	5 x 10	1 of 1
5	GaAs/Ge	4cm x 4cm, 7-mil, WTC	5 x 10	1 of 1
6	GaAs/Ge	4cm x 4cm, 3.5-mil,w/ICG	4 x 4.5	1 of 1
7	InP	2cm x 2cm	4 x 5.5	0 of 1
8	AlGaAs/GaAs	2cm x 2cm, monolithic MBG	3 x 6	0 of 1
9	GaAs/CuInSe$_2$	2cm x 2cm, mech-aligned MBG	6 x 6	0 of 2
10	GaAs	SLATS Concentrator	11 x 13.5	1 of 1
11	GaAs/GaSb	Mini-Dome Fresnel Lens Concentrator, MBG	4.5 x 7.5	1 of 1

BSF = Back-Surface Field
WTC = Wrap-Through Contact
ICG = Integral Coverglass
MBG = Multi-Bandgap

The second solar array in Table I consists of 8 cm x 8 cm silicon cells with wrap-through contacts; these cells are baselined to fly on NASA's Space Station Freedom. The next four arrays (#3 through 6) use the same size GaAs on Ge cells, but differ in electrical or mechanical configuration. The next three arrays (#7 through 9) utilize small (2 cm x 2 cm) new-material cells; the InP cells have high radiation resistance, while the AlGaAs/GaAs and GaAs/CuInSe$_2$ cells are dual junction types having high conversion efficiencies. The final two arrays (#10

663

and 11) are concentrator designs. SLATS, using venetian-blind like metal mirrors, focuses light onto linear strings of GaAs solar cells mounted on the backs of adjacent mirrors (see Figure 1). This design enhances its survivability against man-made threats such as high-powered lasers. The second concentrator (array #11) is the mini-dome fresnel-lens GaAs/GaSb design which has demonstrated extremely high conversion efficiencies through the use of its dual-junction mechanically stacked GaAs and GaSb solar cells and its prismatic coverglass. See Figure 2.
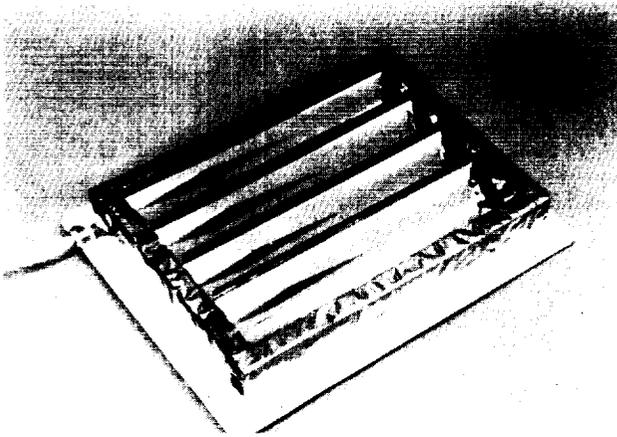
Figure 1. SLATS Concentrator Array.
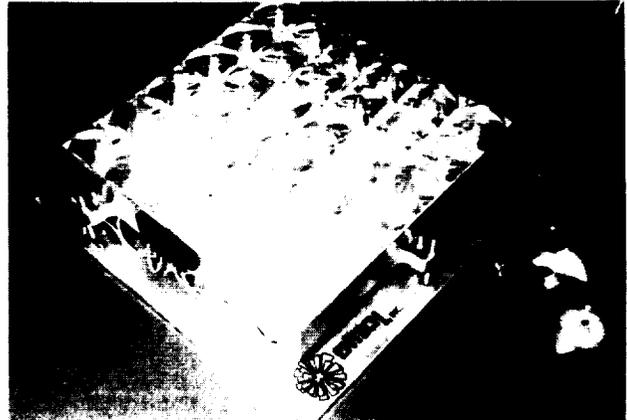
Figure 2. Mini-Dome Fresnel-Lens Dual-Junction GaAs/GaSb Concentrator Array.

The last column of Table I indicates which arrays will be subjected to high-voltage biasing. In some cases, biasing will be applied to only part of an array and not the remaining part. We can thereby investigate long-term radiation damage to array performance for portions (individual "modules") of the array subjected to and not subjected to biasing. High-voltage biasing, besides causing deterioration itself (detectable at the time of the bias measurements), can also increase the susceptibility of the biased array or module to later (or longer-term) contamination or radiation damage. Instrumentation included in PASP Plus will help us to distinguish between different damage effects.

Experiment-Control Instrumentation To simulate a large space-power subsystem operating at high positive or negative voltage levels, PASP Plus instrumentation must provide a multi-step high voltage generator to bias our relatively small test arrays. Some of the eleven arrays will be partitioned into two or three sections, resulting in 16 electrically isolated, individual modules. Ten of the 16 segments will be biased. The high-voltage biasing sequences for each module (one at a time) will consist of four all-positive or all-negative steps (each about 20 sec long) of successively greater voltage levels. The range of bias values (positive or negative) is 50 to 500 Volts. The minimum difference between step values is 10 volts.

In order that positive biasing of a test array can properly simulate actual high-positive-voltage operation of a large array, PASP Plus must provide an additional experiment-control device: an electron emitter. Pegastar's power arrays operate at about +32 Volts with respect to spacecraft ground; however, with respect to the space plasma, the positive side of Pegastar's power arrays will be only a few Volts positive while its frame will float to a potential of around -25 V. Because of the lower mobility of ions, greater highly-negative vehicle surface area is needed to collect the required incoming ion current to balance the incoming current of highly mobile electrons. When (during biasing) a high positive voltage is applied to a small PASP Plus test array, the vehicle frame will be driven further negative, possibly to as much as several hundred Volts depending on array size, bias level, and vehicle metallic surface area (for ion collection). Besides the frame being at a large negative potential, negative-potential contours around the vehicle would limit the accessibility of the plasma's thermal electrons to our positively biased test array except for those directly over it. The incoming-electron acceptance cone for the biased array would be narrowed by the spacecraft's negative-potential contours, artificially lessening the electron current that would otherwise be collected. The plasma "leakage" current measured in our experiment would not be properly representative of the current for a large array operating at a high positive voltage level.

664

If PASP Plus's electron emitter can produce enough <u>outgoing</u> electrons so as to better balance incoming electrons, Pegastar's frame potential will not swing so highly negative. The ensuing negative-potential contours would not be so obtrusive so as to significantly obstruct the plasma "leakage" electrons on their way to our positively biased test array. Therefore, the PASP Plus measurement of leakage current will be more representative (proportionally) of what will happen to large arrays operating at high positive voltage levels.

<u>Interactions-Measuring Instrumentation</u> To measure the characteristics of the arc pulses produced by negative biasing of the experiment's test arrays or array modules, PASP Plus makes use of several electrical transient sensors (ETSs) connected to a Transient Pulse Monitor (TPM). Several E-field sensors at various locations on the surfaces to which the test arrays are mounted are used to measure the parameters of radiated emission from arcs on the test arrays. A current-loop ETS is used to measure arc parameters of emission conducted down the high-voltage power line from the arcing arrays. The pulse characteristics obtained are amplitude, derivative, integral, and number of pulses per time interval.

To measure the plasma "leakage" current produced by positive biasing of the test arrays or array modules, PASP Plus uses an electrometer covering the range of roughly 1 μA to 20 mA. This range should be sufficient to measure from the lowest current of interest up to the largest expected leakage current (i.e., where "snapover" has occurred for our largest array module at the highest bias level [+500 V] and highest plasma density [1×10^6 cm^{-3}] in the nominal APEX mission orbit).

To determine how various environmental interactions affect the arrays, the electrical performance of each of the 16 array modules, whether biased or not, is monitored by taking numerous current-voltage measurements (I-V curves) of the module over the course of mission lifetime. The I-V curves for each module are obtained from the rapid application of dynamically varying resistance values between R = ∞ to R = 0 (corresponding to open-circuit voltage V_{OC} and short-circuit current I_{SC}) to the sun-illuminated array module. Thirty-two digitized measurements of current and voltage are recorded (all within about two sec) for each array module.

<u>Diagnostic Sensors</u> To allow us to investigate fully the linkage between environmental interactions and their performance deteriorating effects, a suite of relevant environmental diagnostics sensors is provided as part of the PASP Plus experiment. These sensors include:
    a. a Langmuir probe (LP) to measure low-energy plasma density and temperature.
    b. an electrostatic analyzer (ESA) to measure 10 eV - 30 keV electron and ion spectra and detect the passage of Pegastar through an auroral region.
    c. an electron/proton radiation dosimeter to measure the high-energy charged particle radiation that damages solar cells, causing the deterioration in array performance measured by the I-V curves. The design of one of the four detection domes of this dosimeter has been specially modified to facilitate the measurement of the 5-10 MeV proton radiation that is an important source of solar-cell degradation (Ref. 2). See Figure 3.
    d. contamination monitors to measure the amount of effluents deposited on array surfaces (leading to decreased sunlight collection and array output power). Contamination sensors will include quartz crystal microbalances (QCMs) and calorimeters. The information from these sensors will allow us to differentiate array performance degradation caused by radiation dosage from that caused by contamination.
    e. a PASP Plus sun incidence-angle sensor to assure us of the alignment of the arrays to the incident solar energy; this alignment is critically important for the concentrator arrays. To meet PASP Plus requirements the Pegastar satellite, using its own sun incidence-angle sensor to provide inputs to the vehicle's magnetic-torque attitude control subsystem, will point its sun-viewing upper-deck honeycomb panel (on which both of our concentrator arrays will be mounted) to within ±0.5° of the sun.

<u>Instrument Layout on Pegastar</u> The various photovoltaic test arrays, the Pegastar and PASP-Plus sun incidence-angle sensors, and the PASP Plus interactions-measuring and diagnostic sensor "heads" separated from their control boxes will be mounted on the hexagon-shaped upper-deck payload shelf and one deployable payload panel, as shown in Figure 4. The other three deployable panels of the satellite will be used for mounting the Pegastar solar arrays used to provide vehicle (and experiment) power. The PASP Plus electronics boxes and those of the other two small experiments (FERRO and CRUX/CREDO) also selected for flight on the APEX mission are mounted on the lower-deck avionics payload panel.

C-4
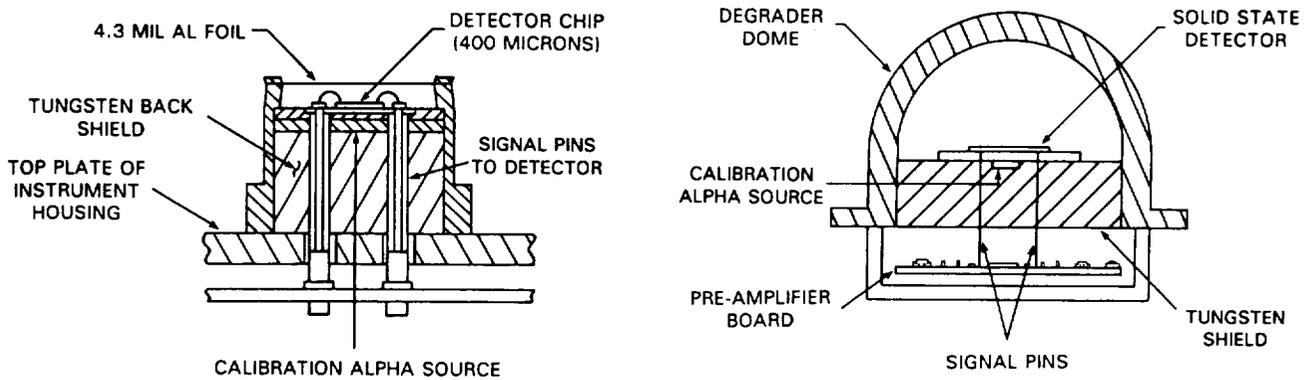
Figure 3. Domes of PASP Plus Dosimeter. Customary design for higher-energy particles is shown at right. Modified design to measure 5-10 MeV protons is shown at left.
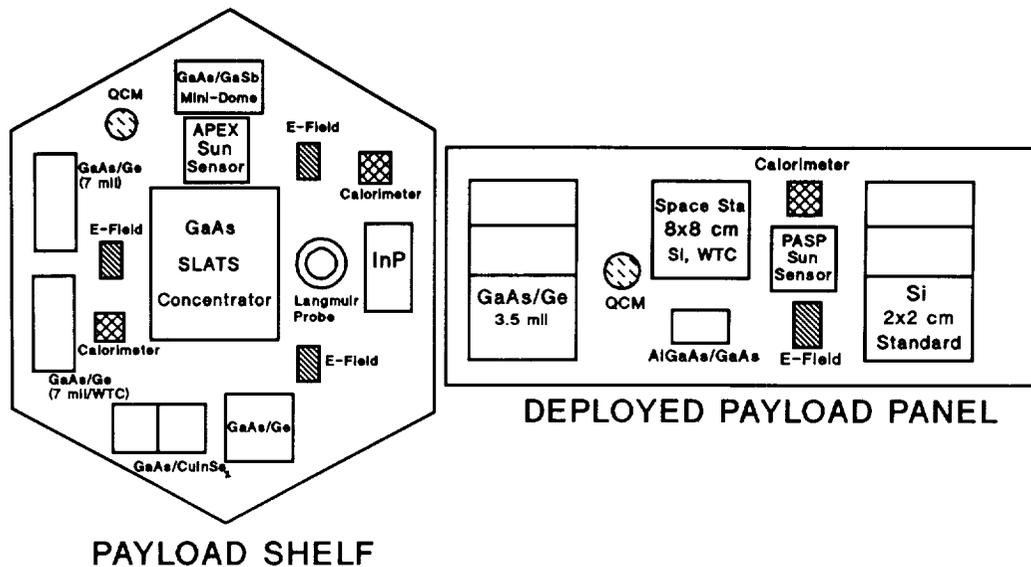
Figure 4. PASP Plus Arrays and Instruments on Pegastar's Upper-Deck Payload Shelf and Deployed Payload Panel.

PEGASTAR/APEX OPERATIONS

The Pegastar satellite characteristics and the expected PASP Plus flight-profile parameters are given in Table II below.

TABLE II. NOMINAL SATELLITE AND FLIGHT-PROFILE PARAMETERS

a. Pegastar Satellite Parameters

SATELLITE BUS: 44 inch diam hexagon
≈60 inch height

WEIGHT TO ORBIT: ≈820 lb

STABILIZATION: three-axis stabilized
sun-pointing to ± 0.5°

ELECTRICAL POWER: 320 - 380 Watts

b. Expected PASP Plus Flight Profile

LAUNCH: Pegasus, Western Test Range,
Late 1992

APOGEE: 1850 km (2000 km preferable)

PERIGEE: 350 km   INCLINATION: 70°

LIFETIME: One to Three Years

666

The high-voltage plasma interactions objectives of PASP Plus will be achieved while Pegastar is near _perigee_ passing through the ionospheric F-region and/or through auroral regions. The radiation degradation objectives of PASP Plus will be achieved while the vehicle is near _apogee_ in the _equatorial_ regions (Pegastar's line of apsides will rotate about 1.5° a day in its orbital plane).

EXPERIMENT STATUS

The functional baseline of the upgraded PASP Plus experiment was established at a meeting of scientists from Phillips Laboratory, Wright Laboratory, Aerospace Corp., Naval Research Laboratory, and NASA Lewis Research Center held at Wright-Patterson AFB in October, 1990. All the basic instruments of the experiment have been designed. Some have been breadboarded; in a few instances, flight-unit construction has begun. In some cases (e.g., the Transient Pulse Monitor and our sun incidence-angle sensor), we have the flight units on hand. Several of the test arrays are on hand; the rest will be delivered by mid-July 1991. Subsequently, they will be tested and mounted on the upper-deck payload shelf and deployable payload panel at the NASA Lewis Research Center. The APEX-mission contractor has begun its design work on the Pegastar satellite; a Preliminary Design Review (PDR) was held in March 1991.

EXPECTED RESULTS

Many investigations of high-voltage interactions have been carried out by the NASA Lewis Research Center, including laboratory and flight-test work (Ref. 3, 4, 5). Various explanations of the causes of arcing from high negative voltage operation have been offered (Ref. 6, 7). The arcing rate (beyond a threshold voltage) appears to be roughly proportional to plasma density, but has a large power-law dependence on voltage level (Ref. 8). For high positive voltage operation, available array power is reduced by electron currents flowing between the array and the surrounding space plasma (Ref. 9). This "leakage" current will depend on the operating voltage, plasma density, exposure of the interconnects, and the geometry of the sheath surrounding the array; various computer simulations have been used to investigate the problem (Ref. 10). Data from the PASP Plus experiment should be very helpful in determining the relationships between various parameters.

APEX's 350 km nominal perigee will provide the greatest electron density (in the order of 10^5 to 10^6 cm^{-3}) at around perigee and allow investigation of space-plasma induced effects over the largest _useful_ range of electron density variations. After several months of flight, a large data base on arc-pulse parameters (negative biasing) and leakage current parameters (positive biasing) as functions of bias levels and types of array will be collected over the mission-achievable ranges of the controlling parameters: plasma density (perigee through apogee), auroral passage, and velocity-vector orientation. This large data base will permit examination of the correlations between all the linkable variables and lead to the establishment of cause-and-effect relationships for high-voltage interactions effects. These relationships will then be available for analytic study, modeling, and code development.

APEX's nominal 70° inclination and 1850 km apogee will allow passage of Pegastar through the lower portion of the inner radiation belt _only_ when apogee occurs near _equatorial_ geomagnetic latitudes. The line of apsides (the perigee-apogee line in the orbital plane) continuously rotates about 1.5° per day throughout the mission. On a long-term basis, Pegastar's apogee will pass through the radiation belt at equatorial latitudes about one-quarter of the time. A higher apogee (up to 2000 or 2200 km) would increase the dosage (or lessen the time to reach a specific dose accumulation), but there are limitations in the Pegasus-Pegastar boosting capability. With some modest improvement in apogee, we expect to obtain sufficient radiation dosage in one year to see array performance degradation in Si cells in the order of 8 to 12 percent. For the more radiation-hard materials (GaAs and especially InP) and the concentrators, the degradation may be only a few percent. Information from the contamination sensors (QCMs and calorimeters) will be used to separate contamination effects from radiation effects. The radiation-induced performance degradation data for all the PASP Plus test arrays will be correlated with the radiation dosage data from our dosimeter to establish cause-and-effect relationships.

Within the first year after a successful PASP Plus flight, correlated PASP Plus data would be made available to the space-power communities in DoD and NASA. Phillips Laboratory, together with Wright Laboratory and NASA's Lewis Research Center, will conduct a series of workshops which will be targeted to major topics of interest such as high-voltage operation and EMI-generation effects. As data on array performance degradation from radiation effects becomes available (1-3 years, depending on flight apogee), additional workshops will be held on

radiation effects on new cell technologies and concentrator arrays. Results from these workshops will be directed towards upgrading space-power design guidelines and test standards.

CONCLUSIONS

Because of the APEX-mission orbit, the upgraded PASP Plus experiment now has the opportunity to investigate both space plasma effects on high-voltage operation and long-term array performance degradation due to space radiation. PASP Plus data on environmental interactions will be given over to the space-system development community so it can capitalize on our results and begin utilizing new-technology photovoltaic arrays on future space systems.

ACKNOWLEDGEMENTS

The support of other Air Force and NASA organizations in the development and future flight of the PASP Plus experiment is gratefully acknowledged. SSD's Space Test Program provides direction and funding of the APEX mission utilizing the Pegastar satellite. Wright Laboratory specified and made arrangements to obtain the PASP Plus test arrays. NASA Lewis Research Center will mount the test arrays on the payload panels and conduct/support testing of the arrays throughout experiment integration.

REFERENCES

1. Guidice, D.A., "Photovoltaic Array Space Power Plus Diagnostics Experiment", Third Annual Workshop on Space Operations, Automation and Robotics (SOAR '89), NASA Conf. Publ. 3059, 1990, pp. 515-519.

2. Tada, H.Y., and Carter, J.R., Solar Cell Radiation Handbook, JPL Publ. 77-56, 1977, Chap. 4.

3. Grier, N.T., "Plasma Interaction Experiment II (PIX II): Laboratory and Flight Results", Spacecraft Environmental Interactions Technology - 1983, NASA CP-2359, AFGL-TR-85-0018, 1985, pp. 333-348.

4. Snyder, D.B., "Discharges on a Negatively Biased Solar Cell Array in a Charged-Particle Environment", Spacecraft Environmental Interactions Technology - 1983, NASA CP-2359, AFGL-TR-85-0018, 1985, pp. 379-388.

5. Ferguson, D.C., "The Voltage Threshold for Arcing Solar Cells in LEO--Flight and Ground Test Results", AIAA 86-0362, AIAA 24th Aerospace Sciences Meeting, Reno, Nevada, Jan. 6-8, 1986.

6. Jongeward, G.A., Katz, I., Mandell, M.J., and Parkes, D.E., "The Role of Unneutralized Surface Ions in Negative Potential Arcing", IEEE Trans. Nucl. Sci., Vol. NS-32, No. 2, 1985, pp. 4087-4091.

7. Hastings, D.E., Weyl, G., and Kaufman, D.,"Threshold Voltage for Arcing on Negatively Biased Solar Arrays", J. of Spacecraft and Rockets, Vol. 27, No. 5, 1990, pp. 539-544.

8. Ferguson, D.C., "Solar Array Arcing in Plasmas", Third Annual Workshop on Space Operations, Automation and Robotics (SOAR '89), NASA Conf. Publ. 3059, 1990, pp. 509-513.

9. Mandell, M.J., Katz, I., Steen, P.G., Schnuelle, G.W., "The Effect of Solar Array Voltage Patterns on Plasma Power Losses", IEEE Trans. Nucl. Sci., Vol. NS-27, No. 6, 1980, pp. 1797-1800.

10. Thiemann, H., and Schunk, R.W., "Particle-in-Cell Simulations of Sheath Formation Around Biased Interconnectors in a Low-Earth-Orbit Plasma", J. of Spacecraft and Rockets, Vol. 27, No. 5, 1990, pp. 554-562.

I

SOLAR CELL ARCING: THE ROLE OF OUTGASSING AND CONTAMINATION

W.J. Marinelli, B.D. Green, B.L. Upschulte, and G. Weyl
Physical Sciences Inc.

D. Hastings
Massachusetts Institute of Technology

E. Aifer
Boston University

ABSTRACT

The effect of outgassing, venting, and thruster firing events on spacecraft system performance has been a long standing issue. Recent laboratory measurements on negatively biased high voltage solar cells at PSI suggest that some currently designed and certainly future space power systems must address/re-evaluate this issue. Our observations show that exposing these cells to moderate levels (10^{-3} torr-min) of H_2O vapor enhances the arcing frequency, while heating to 85°C to remove water vapor significantly reduces the arc frequency. The interaction of the adhesive used to attach the cover glass to the solar cell with ambient water vapor is the key factor in determining arcing rates. Elimination of adhesive exposed to the environmental plasma reduces the arc frequency more than two orders of magnitude, and eliminates any sensitivity to H_2O exposure. The adhesive may also become a source of spacecraft contamination. Macroscopic amounts have been observed to blow off in some arc events, and (we assume) electrostatically precipitate at other negatively biased locations. Data, analysis, and potential impact for future space plateforms will be discussed.

I

Session E2: PLASMA INTERACTIONS - II: SPECIAL USE CODES WITH APPLICATIONS

Session Chair: D. B. Snyder

SPACECRAFT-PLASMA INTERACTION CODES: NASCAP/GEO, NASCAP/LEO, POLAR, DynaPAC, and EPSAT

M. J. Mandell and G. A. Jongeward
Maxwell Laboratories/S-Cubed Division
P. O. Box 1620, La Jolla, CA 92038

D. L. Cooke
Phillips Laboratory/WSSI
Hanscom AFB, MA 01731-5000

Development of a computer code to simulate interactions between the surfaces of a geometrically complex spacecraft and the space plasma environment involves

(1) Defining the relevant physical phenomena and formulating them in appropriate levels of approximation;

(2) Defining a representation for the three-dimensional space external to the spacecraft and a means for defining the spacecraft surface geometry and embedding it in the surrounding space;

(3) Packaging the code so that it is easy and practical to use and to interpret and present the results;

(4) Validating the code by continual comparison with theoretical models, ground test data, and spaceflight experiments.

In this paper we discuss the physical content, geometrical capabilities, and application of five S-CUBED developed spacecraft plasma interaction codes. NASCAP/GEO is used to illustrate the role of electrostatic barrier formation in daylight spacecraft charging. NASCAP/LEO applications to the CHARGE-2 and SPEAR-1 rocket payloads are shown. DynaPAC application to the SPEAR-2 rocket payload is described. EPSAT is illustrated by application to TSS-1, SPEAR-3, and Sundance. The following paper contains a detailed description and application of the POLAR code.

1. INTRODUCTION

S-CUBED has been developing three-dimensional spacecraft-plasma interaction codes since the mid 1970's. During this time there have been great advances in computer hardware, in programming techniques, and in our understanding of the interactions between a spacecraft and its environment. (The latter has been aided by the availability of simulation codes.) Thus, succeeding codes have addressed ever more ambitious goals.

In this paper we briefly describe five S-CUBED developed spacecraft-plasma interaction codes. First, we discuss the physical phenomena we wish to model, their implications for spacecraft operations, and the consequent requirements on computer codes so that they may be used with relative ease and confidence for spacecraft design. Calculations illustrating the types of problems successfully solved will be presented for NASCAP/GEO, NASCAP/LEO, DynaPAC, and EPSAT. In the following paper (Cooke, 1991), David Cooke illustrates in more detail the use of the POLAR code.

2. PHYSICAL PHENOMENA

A uniform plasma satisfies the condition of vanishing space charge - otherwise its electrostatic potential would vary in accordance with Poisson's equation. At a boundary, however, the steady-state condition is to have vanishing net current. It follows that a spacecraft surface perturbs a plasma simply by its presence. Active spacecraft operations, such as applied surface potentials or effluents, serve to enhance that perturbation.

The plasma responds to the presence of a surface by providing currents to it. The incident plasma electrons and ions cause emission of low energy secondary electrons (Katz et al.,1986; Dekker, 1958; Hackenberg and Brauer, 1962; Gibbons, 1966; Dietz and Sheffield, 1975), which add to the photoemission (Feuerbacher and Fiton, 1972) caused by solar ultraviolet light. An insulating surface achieves steady state by reaching a potential at which the various current components are locally in balance. A conducting object (or set of objects) assumes a potential such that the currents balance when integrated over the surface.

The surface-plasma interaction described above is modified by many other effects. Spacecraft orbital motion leads to an inhomogeneous ram-wake structure in the space plasma. The geomagnetic field causes vxB potentials (Lilley et al., 1986) and magnetic insulation (Parker and Murphy, 1967). Plasmadynamic effects result from rapidly applied potentials, from beam operations, or simply from spacecraft caused plasma inhomogeneities. Effluents can be ionized, leading to plasma breakdowns and discharges.

3. CODE OBJECTIVES AND REQUIREMENTS

The objective of a spacecraft-plasma interaction code is to determine self-consistently the potentials and currents on the spacecraft surface and in the surrounding space. The results must be made available in a form which allows assessment of the consequences for spacecraft operations. Operations may be adversely affected by spacecraft charging in the geosynchronous and auroral environments. A high spacecraft floating potential may lead to unacceptable levels of parasitic currents, surface sputtering and contamination, or arcing. The optimal placement and acceptance angle of particle detectors is determined by the structure of the spacecraft sheath.

The first step in a simulation is to create or obtain a representation of the spacecraft surface. A calculation requires not only the geometrical configuration of the spacecraft surface, but also knowledge of the surface materials (in particular, their secondary emission properties) and how surface elements are

672

coupled electrically. A code must be able to model a large volume of surrounding space while maintaining adequate resolution for the phenomena of interest. It must also model a selection of boundary conditions to couple the spacecraft surface to its surrounding volume.

It is highly desirable that use of a computer code spread beyond its authors to the spacecraft science and engineering communities. This requires that thorough documentation be provided on the physical basis of the code as well as on its use. The initial step of creating a spacecraft surface representation must be reasonably convenient. Code input should be straightforward and user-friendly. Graphical and analytical post-processing tools must be available to aid in the interpretation and presentation of results.

To engender confidence in its calculations, the code should be able to reproduce theoretical and experimental results. For three-dimensional codes, the Child-Langmuir planar diode (Child, 1911) and the Langmuir-Blodgett spherical diode (Langmuir and Blodgett, 1924) are standard tests. Ground test experiments involving electron beams and plasma chambers test the code's ability to reproduce relatively well-controlled experimental results. Finally, spaceflight data from well-instrumented vehicles, notably the SCATHA satellite (Stevens and Pike, 1981), the DMSP series of spacecraft (Gussenhoven et al., 1985), and the SPEAR rocket experiments (Katz et al., 1989), validate the ability of the code to simulate plasma interactions under actual space conditions.

Each of the five S-CUBED developed computer codes discussed below was written with a specific primary purpose in mind, but can be used for a variety of aerospace and other applications. NASCAP/GEO (geosynchronous spacecraft charging) (Rubin et al., 1980), NASCAP/LEO (high voltage interactions with dense plasma) (Katz et al., 1981; Mandell et al., 1990a), POLAR (auroral spacecraft charging) (Cooke et al., 1985), and DynaPAC (dynamic plasma interactions) are fully three-dimensional and, among other things, solve some version of Poisson's equation in a large volume of space. EPSAT (Jongeward et al., 1990) differs in that it couples many environment and phenomenological models rather than solving differential equations.

4. NASCAP/GEO

NASCAP/GEO (NASA Charging Analyzer Program for Geosynchronous Earth Orbit) was developed for NASA/Lewis Research Center and the Air Force Geophysics Laboratory during the period 1976-1984. Its primary purpose was to model spacecraft charging in geosynchronous orbit. It has extensive particle tracking modules, including the ability to model electron beam irradiation, which was a popular way to simulate spacecraft charging in terrestrial laboratories.

The low level of plasma screening in the geosynchronous environment imposed the requirement of treating a large volume of empty space. NASCAP/GEO satisfies this requirement by a series of nested grids, each having half the physical dimension of its parent, with the spacecraft contained in the innermost grid. Finite element algorithms were developed to couple the grids with minimal loss of accuracy across the grid boundaries.

A very successful object definition module was written for NASCAP/GEO. Spacecraft were initially defined as collections of rectangular parallelepipeds, wedges, and tetrahedra, which fit comfortably in the cubic grid cells. A simple object definition language and graphical diagnostics made spacecraft surface definition very convenient, even with the crude computer hardware available at the time. Thin

booms, thin plates, and transparent antennas were later added to the original cube-slice object repertoire. Figures 1 and 2 show examples of NASCAP/GEO spacecraft representations.
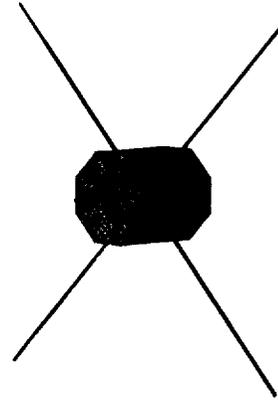
Figure 1. NASCAP/GEO representation of the SCATHA spacecraft.

NASCAP/GEO uses plasma and space charge representations appropriate to the geomagnetic substorm environment. NASCAP/GEO was used extensively to model earthbound electron beam irradiation experiments and flight measurements by the SCATHA spacecraft.
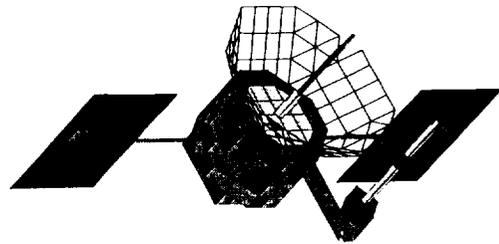
Figure 2. NASCAP/GEO representation of a communications satellite, showing thin plates, long booms, and a transparent antenna.

As an example of a NASCAP/GEO calculation, we present a simulation of a complex effect on a simple object: charging of an insulating sphere in sunlight due to formation of an electrostatic barrier (Mandell et al., 1978). Initially, the sunlit side of the sphere is "grounded" by photoelectron emission, while the dark side gradually charges negative due to incident plasma electrons. Eventually an electrostatic barrier forms to suppress the photoelectron emission, and the entire sphere charges to negative potential.

The NASCAP/GEO representation of a sphere as a twenty-six faceted object is shown in figure 3. Figure 4 shows the NASCAP/GEO produced potential structure around the sphere after the barrier has formed. The sun is incident from the upper right, and the potential field exhibits a saddle structure in front of the sunlit surface. Particle trajectories (figure 5), calculated using NASCAP/GEO's "DETECTOR" module, illustrate the effect of the barrier in blocking all but the highest energy photoelectrons.

Following barrier formation, the sunlit surface follows the dark surface in achieving high negative potential. As shown in figure 6, the process is characterized by a long time constant of

673

many minutes. This process has been observed on actual spacecraft (Olsen and Purvis, 1983).
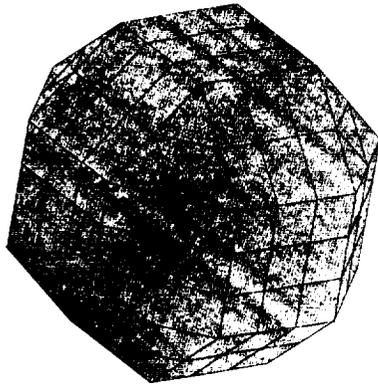
Figure 3. NASCAP/GEO representation of a sphere as a twenty-six faceted object.
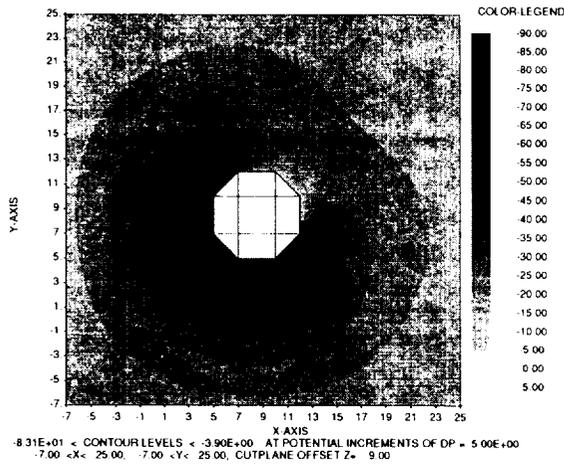
Figure 4. Potential structure around the sphere following barrier formation. Sunlight is incident from the upper right. The shaded surfaces are at -83 volts, the least negative surface potential is -4 volts, and the saddle point is at approximately -15 volts.
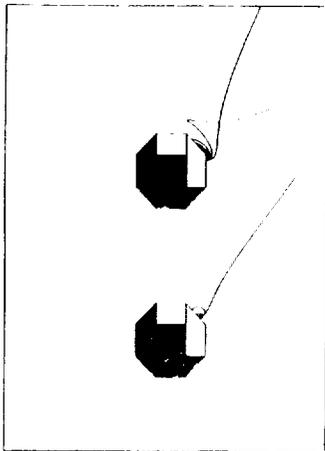
Figure 5. Particle trajectories for emitted photoelectrons of various energies. Only the most energetic electrons can escape over the electrostatic barrier.
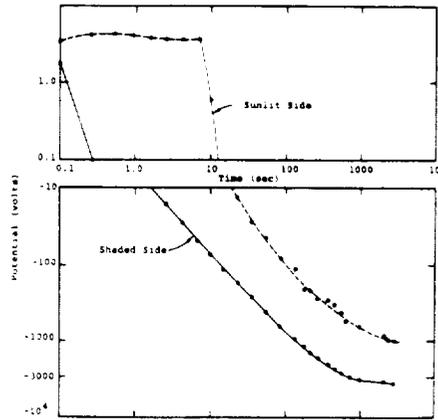
Figure 6. Time dependence of surface potentials for daylight charging.

5. NASCAP/LEO

NASCAP/LEO (Katz et al., 1981; Mandell et al., 1990a), was developed for NASA/Lewis Research Center during the period 1980-1990 for the purpose of studying plasma interactions in the dense, short Debye length plasma characterizing low- Earth orbit. It pioneered methods for embedding an arbitrary object (defined by industry-standard finite element pre-processors) in a cubic grid, with local subdivision for the resolution of small but important object features. It incorporates models for solar array surfaces, solar array circuitry, and hydrodynamic ion expansion. It has been applied to many plasma chamber experiments and rocket flights.

To avoid the need for obtaining space charge from particle simulations, NASCAP/LEO pioneered an analytic representation for space charge as a function of the local potential and electric field. The formula

$$\rho / \varepsilon_0 = -\left(\phi / \lambda_D^2\right)\left(1 + |\phi / \theta| C(\phi, E)\right) / \left(1 + (4\pi)^{1/2} |\phi / \theta|^{3/2}\right)$$

$$C(\phi, E) = |\theta / \phi| \left[\left(R_{sh} / r\right)^2 - 1\right]$$

$$\left(R_{sh} / r\right)^2 = 2.29 |E\lambda_D / \theta|^{1.262} |\theta / \phi|^{.509}$$

reduces to linear screening for low potentials, and contains acceleration and convergence effects to reproduce a Langmuir-Blodgett spherical diode at high potentials.

NASCAP/LEO was used to perform a series of calculations for the CHARGE-2 rocket (Mandell et al., 1990b; Neubert et al., 1990). The CHARGE-2 mother vehicle could be held at a known negative potential by biasing it relative to the daughter, or taken to positive potential by electron beam emission. The electron and ion currents collected by the mother vehicle could be determined by measuring the tether current. The measurements compared well with NASCAP/LEO simulations.

In this paper we discuss performance of the "floating probes," which were intended to measure the vehicle potential and the sheath profile. Figure 7 shows the NASCAP/LEO model of CHARGE-2. (Note that, unlike a NASCAP/GEO object, the shape need not conform to the cubic grid.) The floating probes are seen extending up to one meter from the rocket surface. Under negative bias conditions, the potential of a floating probe is determined by equating the collected plasma ion current to the current through the 100 MΩ probe impedance.

674
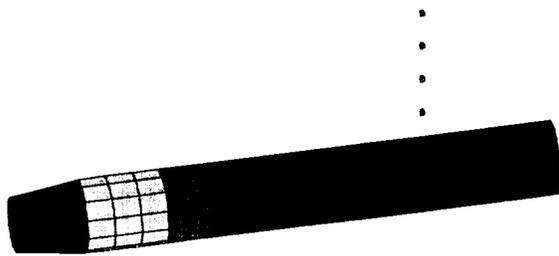
Figure 7. NASCAP/LEO model of CHARGE-2, showing floating probes.
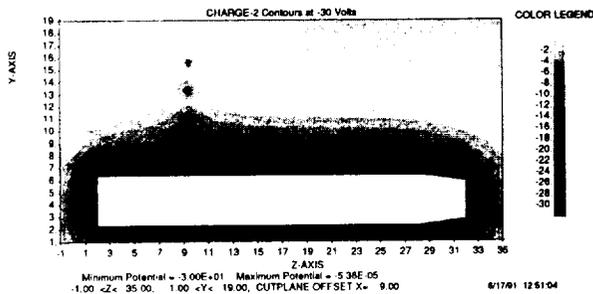
Figure 8. Potential around CHARGE-2 mother vehicle at -30 volts, showing perturbation by floating probes.
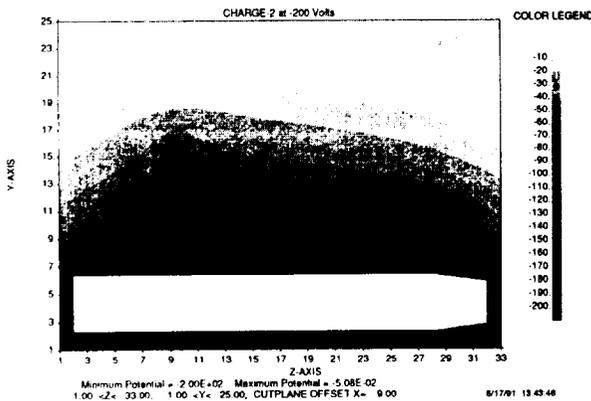
Figure 9. Potential around CHARGE-2 mother vehicle at -200 volts, showing floating probes inside the sheath.

Figure 8 shows the self-consistent potential structure with the mother vehicle at -30 volts. The probes are seen to produce a substantial sheath perturbation. Even though outside the sheath, the outermost probe cannot collect enough plasma ion current to provide an accurate potential measurement. Figure 9 shows the self-consistent potential structure with the mother vehicle at -200 volts. For this case, all of the probes are inside the sheath, and the 100 MΩ probe impedance essentially grounds the probes to the spacecraft. Figure 10 shows the calculated and experimental results for the spacecraft potential measured by the outermost probe as a function of the actual spacecraft potential (i.e., the tether bias). The floating probe measures about 70% of the actual potential up to about 20 volts. At higher spacecraft potentials, when the outermost probe is within the sheath, the measurement fails completely.

NASCAP/LEO played a crucial role in modeling the SPEAR-1 rocket experiment (Katz et al., 1989). SPEAR-1 (figure 11) contained two eight-inch diameter spheres which could be positively biased to a total of 45 kilovolts. A hollow cathode plasma contactor was intended to control the rocket body
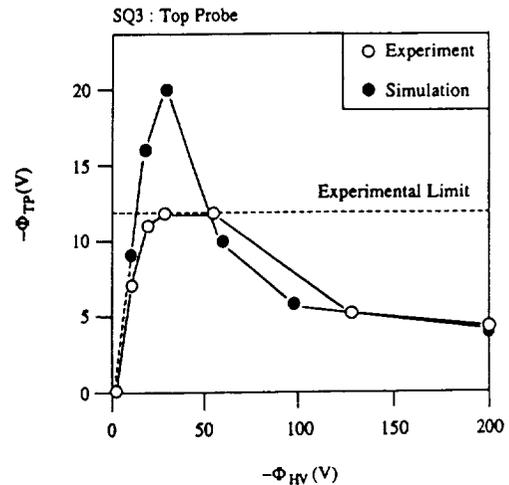
Figure 10. Experimental and NASCAP/LEO simulation results for the floating probe potential measurement vs. actual mother vehicle potential.
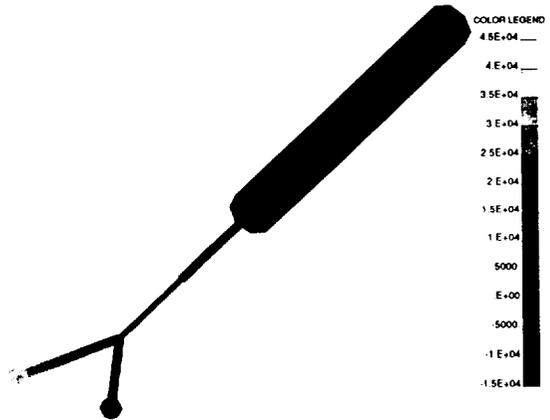
Figure 11. NASCAP/LEO model of the SPEAR-I rocket payload.

potential. With the failure of the hollow cathode door, this high voltage experiment became asymmetric and fully three-dimensional. The overlapping, bipolar sheath structure (figure 12) modified both ion collection by the rocket and magnetically limited electron collection by the sphere. NASCAP/LEO was used to calculate the electron and ion currents as a function of rocket potential, and thus determine the floating potential of the experiment.

Figure 13 shows a representative electron trajectory. If the sheath contours were concentric with the spherical probe, such an electron would ExB drift around a given potential contour and never be collected. In the actual case, however, the electron ExB drifts into a high field region and is collected by the sphere.

6. POLAR

The POLAR code (Cooke et al., 1985; Cooke, 1991) was developed for the Air Force Geophysics Laboratory during the time NASCAP/LEO was being written for NASA. POLAR was tuned for the auroral environment and for wake simulation. POLAR adopted an object definition language from NASCAP/GEO, and a dense plasma formulation from NASCAP/LEO. A "bread-slice" grid scheme was developed
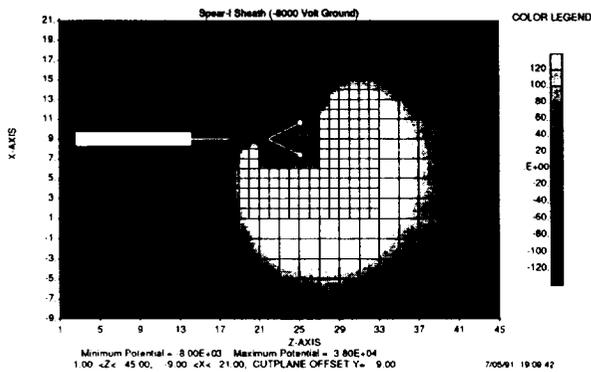
Figure 12. Bipolar sheath structure about SPEAR-I. (Sphere bias 45 kV, rocket potential -8 kV.)

$n_e = n_i = 3 \times 10^{10} \, m^{-3}$

$\theta_e = \theta_i = 0.1 \, eV$

$B = 0.4 \, gauss$

$\phi_{sphere \, 1} = 45 \, kV$

$\phi_{sphere \, 2} = 0 \, kV$

$\phi_{ground} = -6 \, kV$
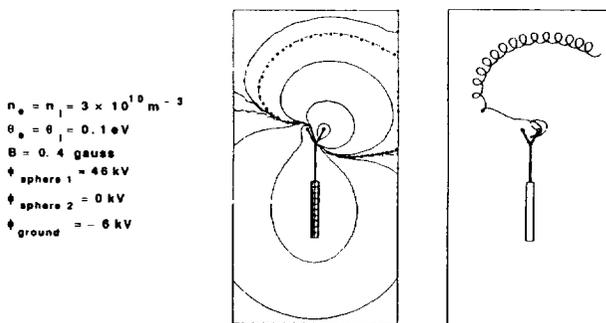
Figure 13. Potential contours and representative electron trajectory for SPEAR-1, showing failure of magnetic insulation due to sheath distortion.

so that the computational space could extend arbitrarily far into the wake of a spacecraft. The complex auroral charging environment is represented by a Fontheim fit (Fontheim et al., 1982). POLAR is able to correct the initially analytic space charge with particle trajectory data to achieve a self-consistent result, even in the presence of magnetic fields (Mandell et al., 1990b). In the following paper, David Cooke discusses more details about the POLAR code and its application to the DMSP satellites.

7. DynaPAC

DynaPAC, under development since 1989 for Geophysics Laboratory, Phillips Laboratory/WSSI, and SDIO, represents the next generation of three-dimensional code. It is being written for the computer hardware and the software techniques of the nineties. While it will be able to reproduce the results of the earlier codes, its primary focus is on dynamic effects for high power space applications. DynaPAC models space with arbitrarily nested cubic grids. Its high-order cubic elements provide more accurate potentials than the earlier codes and strictly continuous electric fields for particle tracking. Object definition, adapted from the NASCAP/LEO code, interfaces to industry-standard finite element preprocessors. Interactive screen interfaces make it easy to generate input for the DynaPAC modules. It is planned to include a selection of algorithms for various calculational tasks, so that a user can choose that most appropriate for his application, or use DynaPAC as a workbench for testing new algorithms. A programmer-friendly database language simplifies postprocessing tasks and construction of interfaces to other codes.

DynaPAC was used to model the SPEAR-2 payload tests in a very large vacuum chamber in PlumBrook, Ohio. SPEAR-2 contained a pulsed power system with voltages up to -100 kV and pulse widths up to 50 microseconds. By calculating the incident plasma currents to the high voltage components, DynaPAC predicted the observed anomalous measurement of the transformer secondary voltage in the presence of plasma.

The DynaPAC geometric model of SPEAR-2 (figure 14) clearly shows the high voltage components, including the pulse transformer, the klystrode battery pack, and, most prominently, the voltage divider probe used to measure the transformer secondary voltage. To calculate currents to the components, space is initially with a uniform ion distribution. DynaPAC then alternately solves for the space potentials and moves ions in those potentials, with the payload voltage following the prescribed risetime. By three microseconds (figure 15), ions are seen to be drained or expelled from high field regions, and to be converging on the voltage divider. The incident ion currents to the high voltage components (figure 16) are an order of magnitude higher than the equilibrium currents, which can be calculated either by NASCAP/LEO or by DynaPAC. Most of the incident ion current is to the high voltage end of the voltage divider probe. After adding the effects of ion- generated secondary electrons and evaluating the effect of the total injected current on the voltage divider circuit, we predicted that these currents would cause the transformer secondary voltage measurement to be low by about a factor of two. Experimentally, the measurement functioned nominally in vacuum, but was a factor of two low in plasma in agreement with the prediction.
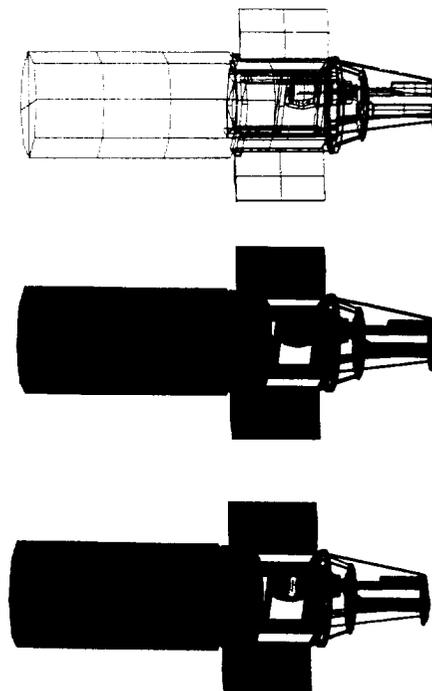
Figure 14. DynaPAC geometric model of SPEAR-2, showing wire-frame model (top), component-coded model (center), and potential-coded model (bottom).
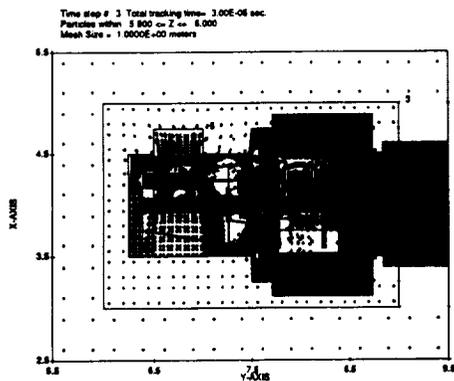
Figure 15. Ion macroparticle positions three microseconds into a SPEAR-2 high voltage pulse.
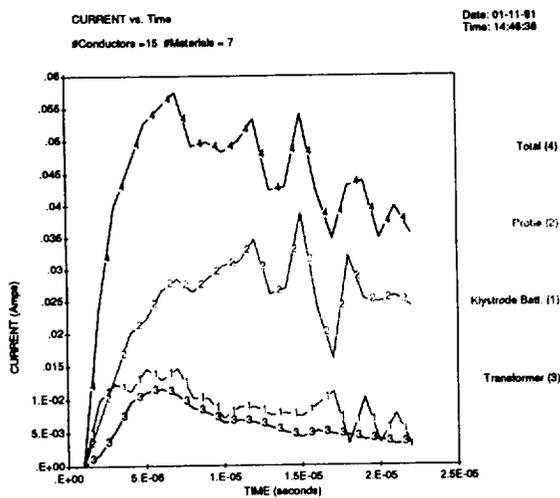
Figure 16. Time-dependent incident ion currents to the SPEAR-2 high voltage components.

8. EPSAT

EPSAT (Jongeward et al., 1990), under development since 1989 for NASA/Lewis Research Center and SDIO, is a very different type of code. Rather than solving three-dimensional equations, EPSAT uses simple, approximate analytic models to evaluate plasma interactions. An intelligent database couples a large number of interaction and environment models so that parameter studies are easily and efficiently performed.

EPSAT integrates quick running models into a unified desktop analysis tool (figure 17). The user has access to individual model results, such as the neutral environment at a point, or to coupled multi-step analyses, such as the total fluence of oxygen. EPSAT currently includes over one hundred models, many of which are shown in figure 18. The selection of models was based on the natural environment design needs for SDIO power systems. Desktop analysis tools for different applications can be constructed by appropriate addition or replacement of modules. A current project is development of an Environmental WorkBench for Space Station Freedom based on the EPSAT architecture.
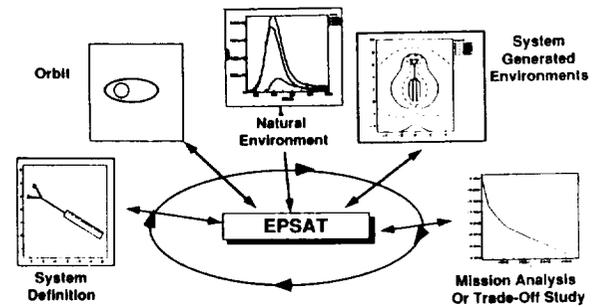
Figure 17. EPSAT integrates quick running models for orbits, environments, and interactions into a comprehensive desktop tool.
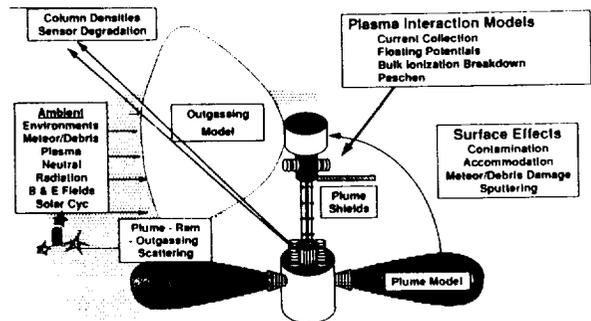
Figure 18. EPSAT has over one hundred models selected, based on natural environment design needs for SDIO power systems.
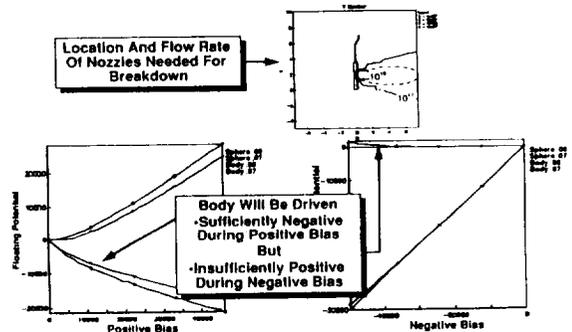
Figure 19. EPSAT was used to estimate floating potentials and effluent densities in the conceptual design of SPEAR-3.

EPSAT was used in the conceptual design of SPEAR-3. (See figure 19.) SPEAR-3 provides a test bed for spacecraft neutralization techniques. Rocket potential is provided by applying a high voltage to a SPEAR-1 type probe. EPSAT predicted that interesting negative potentials should be readily attained, but that only fairly small positive potentials could be achieved on the rocket. EPSAT was also used to predict effluent density profiles from gas release experiments.

At Phillips Laboratory (Boston) EPSAT was applied (M. Oberhardt, private communication) to the TSS-1 mission. (See figure 20.) High subsatellite potentials were predicted. Because only simplified analytic calculations are done, EPSAT can obtain results for many times during a mission, retrieving required parameters, such as plasma density or magnetic field, from accepted environment models. The floating potential plot reflects the night-day variation of plasma density and the diurnal magnetic field variation as the earth turns under the

677

orbit, and predicts a large variation in subsatellite potential. Also, the neutral density due to gas leaks appears to be high enough to lead to the possibility of breakdown.

Phillips Laboratory (Albuquerque) (B. Lillie, private communication) used EPSAT to examine the return flux of ACS effluent to optical sensors on the Sundance satellite. (See figure 21.)
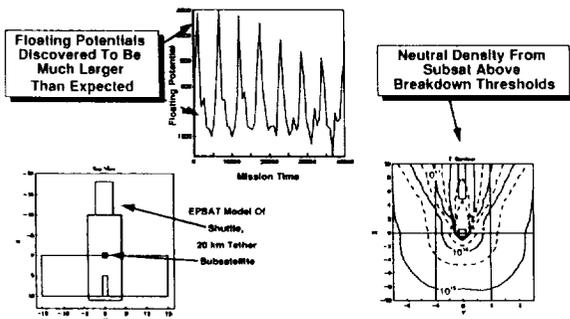
Figure 20. EPSAT was applied to the TSS-1 subsatellite to calculate orbital and diurnal potential variations and effluent densities.
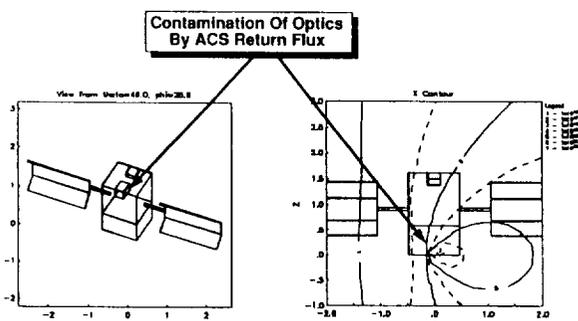
Figure 21. EPSAT was used to examine return flux of ACS effluent to optical sensors on the Sundance satellite.

9. CONCLUSION

It is practical and useful to simulate spacecraft-plasma interactions with three-dimensional, realistic geometry. The results are useful both for spacecraft design and for understanding and analysis of spaceflight data. Knowledge gained in the development and validation of three- dimensional codes has resulted in simplified interaction models which can be integrated with orbital and environment models in comprehensive design tools such as EPSAT. Appropriate simulations should be performed throughout the design phase of a spacecraft or space experiment to assure that no unexpected or harmful effects will occur due to interaction with the space environment.

REFERENCES

Child, C. D., "Discharge From Hot CaO," PHYSICAL REVIEW, 32, 492, 1911.

Cooke, D. L., "Validation and Applications of POLAR Code," 1991.

Cooke, D. L., Katz, I., Mandell, M. J., Lilley, Jr., J. R., and Rubin, A. J., "Three-Dimensional Calculation of Shuttle Charging in Polar Orbit," SPACECRAFT ENVIRONMENTAL INTERACTIONS TECHNOLOGY-1983, NASA CP2359, AFGL-TR-85-0018, 1985.

Dekker, A. J., "Secondary Electron Emission," SOLID STATE PHYSICS, 6, 1958, 251.

Dietz, L. A. and Sheffield, J. C., "Secondary Electron Emission Induced by 5-30-kev Monatomic Ions Striking Thin Oxide Films," JOURNAL OF APPLIED PHYSICS, 46, 4361-4370, 1975.

Feuerbacher, B., and Fiton, B., "Experimental Investigation of Photoemission from Satellite Surface Materials," JOURNAL OF APPLIED PHYSICS, 43, 1972, 1563.

Fontheim, E. G., Stasiewicz, K., Chandler, M. O., One, R. S. B., Gombosi, E., and Hoffman, R. A., "Statistical Study of Precipitating Electrons," JOURNAL OF GEOPHYSICAL RESEARCH, 87, A5, 1982.

Gibbons, D. J., "Secondary Electron Emission," HANDBOOK OF VACUUM PHYSICS, 2, 1966, 299.

Gussenhoven, M. S., Hardy, D. A., Rich, F., Burke, W. J., and Yeh, H.-C., "High-Level Spacecraft Charging in the Low-Altitude Polar Auroral Environment," JOURNAL OF GEOPHYSICAL RESEARCH, 90, 1985, 11009.

Hackenberg, O. and Brauer, W., Advances in Electronics and Electron Physics, 16, 1962, 145.

Jongeward, G. A., Kuharski, R. A., Kennedy, E. M., Wilcox, K. G., Stevens, N. J., Putnam, R. M., and Roche, J. C., "The Environment Power System Analysis Tool Development Program," CURRENT COLLECTION FROM SPACE PLASMAS, NASA CP3089, 1990.

Katz, I., Jongeward, G. A., Davis, V. A., Mandell, M. J., Kuharski, R. A., Lilley, Jr., J. R., Raitt, W. J., Cooke, D. L., Torbert, R. B., Larson, G., and Rau, D., "Structure of the Bipolar Plasma Sheath Generated By SPEAR I," JOURNAL OF GEOPHYSICAL RESEARCH, 94, 1989, 1450.

Katz, I., Mandell, M. J., Schnuelle, G. W., Parks, D. E., and Steen, P. G., "Plasma Collection by High-Voltage Spacecraft in Low Earth Orbit," JOURNAL OF SPACECRAFT AND ROCKETS, 18, 1981, 79.

Langmuir, I., and Blodgett, K. B. "Current Limited By Space Charge Flow Between Concentric Spheres," PHYSICS REVIEW, 24, 1924, 49.

Lilley, J. R. Jr., Katz, I., and Cooke, D. L., "V x B and Density Gradient Electric Fields Measured From Spacecraft," JOURNAL OF SPACECRAFT AND ROCKETS, 23, 1986, 656.

Mandell, M. J., Katz, I., Davis, V. A., and Kuharski, R. A., "NASCAP/LEO Calculations of Current Collection," CURRENT COLLECTION FROM SPACE PLASMAS, NASA CP3089, 1990a.

Mandell, M. J., Katz, I., Schnuelle, G. W., Steen, P. G., and Roche, J. C., "The Decrease in Effective Photocurrents due to Saddle Points in Electrostatic Potentials near Differentially Charged Spacecraft," IEEE TRANSACTIONS ON NUCLEAR SCIENCE, NS-25, 1978, 1313.

Mandell, M. J., Lilley, Jr., J. R., Katz, I., Neubert, T., and Myers, N. B., "Computer Modeling of Current Collection By the CHARGE-2 Mother Payload," GEOPHYSICAL RESEARCH LETTERS, 17, 1990b, 135.

Neubert, T., Mandell, M. J., Sasaki, S., Gilchrist, B. E., Banks, P. M., Williamson, P. R., Raitt, W. J., Meyers, N. B., Oyama, K. I., and Katz, I., "The Sheath Structure Around a Negatively Charged Rocket Payload," JOURNAL OF GEOPHYSICAL RESEARCH 95, 1990, 6155.

Olsen, R. C., and Purvis, C. K., "Observations of Charging Dynamics," JOURNAL OF GEOPHYSICAL RESEARCH 88, 1983, 5657.

Parker, L. W., and Murphy, B. L., "Potential Buildup on An Electron-Emitting Ionospheric Satellite," JOURNAL OF GEOPHYSICAL RESEARCH, 72, 1967, 1631.

Rubin, A. G., Katz, I., Mandell, M. J., Schnuelle, G., Steen, P., Parks, D., Cassidy, J. J., and Roche, J. C., "A Three-Dimensional Spacecraft-Charging Computer Code," SPACE SYSTEMS AND THEIR INTERACTIONS WITH EARTH'S SPACE ENVIRONMENT, H. B. Garrett, and C. P. Pike, ed., Progress in Astronautics and Aeronautics, Vol. 71, AIAA, 1980.

Stevens, N. J. and Pike, C. P., SPACECRAFT CHARGING TECHNOLOGY, 1980, NASA CP2182, 1981.

A COMPARISON OF TWO PLASMA MODELS

Russell Cottam

Sverdrup Technology
2001 Aerospace Parkway
Brookpark, Ohio 44142

Abstract

The time dependent behavior of a plasma which surrounds a highly biased conducting sphere is considered. The plasma is treated as either a cold two component fluid or as a warm plasma whose time development can be found by solving the Vlasov equation. Both models demonstrate oscillatory behavior, but the electric fields predicted by the models are quantitively quite different in regions close to the surface of the sphere and very similar otherwise. A broadening of the electron distribution function with time is observed indicating local heating of the plasma near the surface of the sphere.

Introduction

A problem of particular interest at a time when the production of high powered space systems is imminent is the determination of how surfaces which are biased at a potential well above or below plasma ground interact with the ambient plasma. An example of such a problem is the analysis of the return current from a plasma after the surface potential of a structure in space has been violently changed by an arc, which is essentially a short circuit to plasma ground. In seeking to understand the dynamic behavior of space plasmas what is needed is a description of the average behavior that emerges when a sea of charged particles interacts electromagnetically with itself and with any object immersed in it. A mathematically succinct if numerically unsolvable set of equations exists which captures the dynamics of plasma interactions in regions of low density where two particle correlations may be safely ignored. These are the Vlasov-Maxwell equations. In the case presently under consideration in which no magnetic field exists these equations are:

$$\frac{\partial f_s(\vec{r},\vec{v},t)}{\partial t} = -\vec{v}\cdot\vec{\nabla}f_s(\vec{r},\vec{v},t) - \vec{a}_s\cdot\vec{\nabla}_v f_s(\vec{r},\vec{v},t) \tag{1}$$

$$\vec{a}_s = \frac{q_s}{m_s}\vec{E} \tag{2}$$

$$\frac{\partial \vec{E}(\vec{r},t)}{\partial t} = -\frac{\Sigma \vec{j}_s(\vec{r},t)}{\epsilon} \tag{3}$$

$$\vec{j}_s(\vec{r},t) = q_s \int \vec{v} f_s(\vec{r},\vec{v},t) d^3 v \tag{4}$$

In the above s refers to any plasma species. (1) is the Vlasov equation and (3) is Ampere's law which gives the time dependence of the electric field in the absence of magnetic fields. Ampere's law is coupled to the Vlasov equation through (4) which gives the current carried by the plasma constituents. The distribution function f(r,v,t) gives the probability of finding a plasma particle in an infinitesimal volume of phase space. Knowledge of this function permits one to calculate any characteristic of the plasma. For instance, the integration of f(r,v,t) over velocity space gives the plasma density, and the integration of the product of f(r,v,t) and velocity over velocity space gives the currents in the plasma (ref. 1). The Vlasov equation is unfortunately a seven dimensional quantity which makes finding its solution in the general case numerically intractable, (however, see ref. 2).

A simplification occurs when it is recognized that only quantities averaged over velocity space are of interest. These averages are found by calculating the moments of the Vlasov equation. The first two moment equations are:

$$\frac{\partial \rho(\vec{r},t)}{\partial t} = -\vec{\nabla}\cdot(\rho \vec{v}) \tag{5}$$

$$\frac{\partial \vec{v}}{\partial t} = -\vec{v}\cdot\vec{\nabla}\vec{v} + \frac{q}{m}(\vec{E}) - \frac{\vec{\nabla}\cdot P}{\rho}\frac{q}{m} \tag{6}$$

P is the pressure tensor and   is the charge density. To complete the description a complete set of moment equations is neccessary to predict the time evolution of the pressure tensor and related quantities. If the set of moment equations can be truncated then a fluid description of plasma behavior is possible (ref. 1,3,4). Otherwise one must solve the more fundamental coupled Vlasov-Maxwell equations themselves. In this paper both of these approaches will be used to look at the relatively simple case of a highly biased conducting sphere immersed in a plasma with no magnetic fields present.

The Cold Plasma

In the case where the electrostatic field energy is much larger than the plasma thermal energy it is reasonable to treat the plasma as a cold plasma. In this case the pressure tensor vanishes and equations (5) and (6) form a self contained set of equations if the electric field is known. The electric field is found by solving (3) with:

$$\vec{j}_s = \rho_s \vec{v}_s \tag{7}$$

Equations (3),(5),(6), and (7) form the numerical basis of the cold plasma model. Unfortunately in their present differential form these equations are very difficult to use near charged surfaces. One plasma species or another will certainly be repelled from such a surface and there will be a region, however thin, where the density of this species will become vanishingly small. Equation (6) decouples the velocity from the density of each species. To solve this equation near a surface will require the assignment of an average velocity to a region of space in which there are no particles. The arbitrariness of this assignment will be propagated into regions where particles exist and the calculation will lose accuracy, perhaps catastrophically. To avoid this dilemma a control volume approach is used. The integrals:

$$4\pi\frac{\partial\int\rho(\vec{r},t)r^2dr}{\partial t} = -\int_{surface}\rho(\vec{r},t)\vec{v}\cdot d\vec{a} \tag{8}$$

$$4\pi\frac{\partial\int\rho\vec{v}r^2dr}{\partial t} = -\int_{surface}\rho\vec{v}\vec{v}\cdot d\vec{a} + 4\pi\int\rho\vec{a}r^2dr \tag{9}$$

give the change in charge density due to charge flow into and out of a region, and the change in current density due to the current flow into and out of a region and due to accelerations which occur within that region. Equations (8) and (9) are evaluated within and on concentric spherical shells. This results in:

$$\frac{\Delta\rho r^2}{\Delta t} = \frac{\rho r_i^2 v_i - \rho_o r_o^2 v_o}{\Delta r} \tag{10}$$

$$\frac{\Delta\rho r^2 v}{\Delta t} = \frac{\rho r_i^2 v_i^2 - \rho_o r_o^2 v_o^2}{\Delta r} + \rho r^2 a \tag{11}$$

(i and o refer to the quantities along the inner and outer radii respectively, of a given
shell). The current density found by solving equation (11) is divided by the charge density (eqn. 10) to give the average velocity.

First order upwind differencing is used to evaluate these difference equations with special attention paid to regions where velocity reversal occurs to ensure that the differencing scheme is always conservative (ref. 5,6). The calculations were done in Fortran on a Mac IICX with an Apple/Unix attachment.

One of the interesting features of the cold plasma model is displayed in figure 1. This graph shows the current collected on the surface of a 1 meter sphere which was initially biased at +140 volts with respect to a neutral hydrogen plasma. The surface charges by attracting electrons to itself. After awhile, the rate of charging begins to decrease as the accumulating electrons begin to repel those that follow behind. Soon the induced field is strong enough to repel all approaching electrons and the repelled electrons expand outward until the cloud of ions left behind begins to pull them back again. The swarm of re-attracted electrons strikes the surface, increasing its negative charge and turning back further incoming electrons. These repelled electrons stream out into space again and the process repeats itself at a rate determined by the inverse of the electron plasma frequency.

The Warm Plasma

There is no angular dependence in any quantity under consideration. This greatly reduces the complexity of the Vlasov equation and it becomes amenable to numerical solution. In spherical coordinates:

$$f(\vec{r},\vec{v},t) = f_\theta(v_\theta)f_\phi(v_\phi)f_r(r,v_r,t) \tag{12}$$

$$\frac{\partial f_\theta(v_\theta)}{\partial t} - \frac{\partial f_\phi(v_\phi)}{\partial t} - 0 \tag{13}$$

$$\frac{\partial f_r(r,v_r,t)}{\partial t} - -v_r \frac{\partial f_r(r,v_r,t)}{\partial r} - a_r \frac{\partial f_r(r,v_r,t)}{\partial v_r} \tag{14}$$

Equation (1) is now a one dimensional equation and the coupled set of equations (1), (2), (3), (4) may now be solved numerically using upwind differencing. The strategy is as follows:

The electric field at time step n is used to find the distribution function at time step n+1. This is used in equation (4) to find the current density at time step n+1. This current is then used in equation (3) to find the electric field at time step n+1, (backward differencing), and the process is repeated to project the solution from time step n+1 to time step n+2. The boundary values at the surface of the sphere and at the last grid point are calculated by linear extrapolation from the two neighboring grid points, except that at the surface of the sphere there are assumed to be no emitted or backscattered plasma particles and the distribution function vanishes for all positive values of radial velocity.

The results of this calculation can be seen in figures (2-5), and (6-9). Figures (2-5) indicate that the warm and cold plasma models agree in their predicted electric values only at points at least 2.5 meters from the boundary of the sphere. Within that distance the fields are out of phase and the warm plasma model fields have less amplitude and are more rounded. The time evolution of the electron probability distribution function at the tenth grid point is shown in figures (6-9). They reveal two interesting characteristics. The distribution function does not retain its Gaussian shape, becoming peaked alternately on its left and then its right hand side, and once again oscillations at about the electron plasma frequency occur.

Conclusions

Two conclusions are self evident. The warm and cold plasma models give dissimilar results near the surface of the sphere even when the initial electric field is very strong. Figure 10 indicates that some of the electrostatic field energy initially available may be converted into thermal energy as the electrons flow back and forth in the the field induced by their own motion. This possibility is not allowed in a cold plasma model. The oscillatory nature of the solutions should also be noted. With no damping mechanism in either model no steady state solution exists, and a longitudinal electric field which oscillates at the electron plasma frequency is set up. This is in agreement with the results predicted by a linearized cold fluid model.

The Vlasov equation is a tantilizing object. Its solution gives complete information about the plasma including its internal pressure gradients and currents. That solution, however, can only be found in a limited number of situations. But one such case occurs when a longitudinal electric field is established around a simple object, and understanding the interactions between such fields and the plasma is necessary for understanding spacecraft charging.

References

(1) Montgomery, D.C. & Tidman, D.A., PLASMA KINETIC THEORY
McGraw-Hill, N.Y., N.Y. 1964 p. 11, 194

(2) Johnson, R.L., & Pal, S.K., "The Numerical Solution of Hyperbolic Systems
Using Bicharacteristics", MATHEMATICS OF COMPUTATION, Vol. 26, #118,
April 1972, pp. 377-392

(3) Holt, E. H., & Haskell, R.E., FOUNDATIONS OF PLASMA DYNAMICS
     Macmillan, N.Y., N.Y. 1965, p. 176

(4) Hockney, R.W., & Eastwood, J.W., COMPUTER SIMULATIONS USING PARTICLES
     Adam Hilger, Bristol and New York, 1988 p. 16

(5) Roache, Patrick, COMPUTATIONAL FLUID DYNAMICS
     Hermosa Publishers, Albuquerque, N.M., 1985, p. 73

(6) Tajima Toshiki, COMPUTATIONAL PLASMA PHYSICS
     Addison Wesley, Redwood City California, 1989, p. 177

I

CURRENT COLLECTED AT SURFACE OF SPHERE

sphere is 1 meter in radius, initially biased
to +140 volts with respect to plasma potential.
plasma density is 1.0 e+5 particles/cm**3. temp. =.1 ev
charging spikes separated by 1/(electron plasma frequency)

fig. 1

685
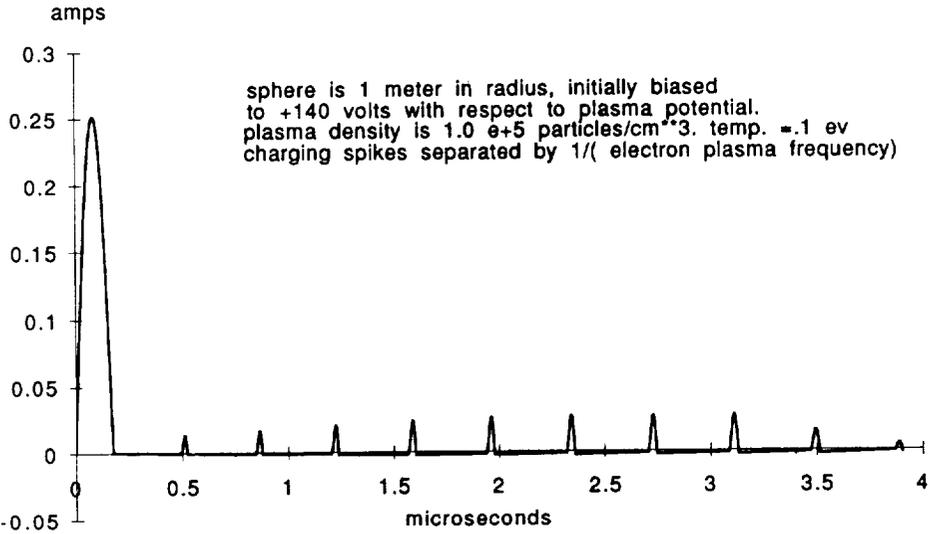
ELECTRIC FIELD COMPARISONS

time = .1 microseconds

cold plasma
warm plasma

fig. 2

time = 1.1 microseconds

warm plasma
cold plasma

fig. 3

time = 2.1 microseconds

cold plasma
warm plasma

fig. 4

volts/meter

meters

cold plasma
warm plasma
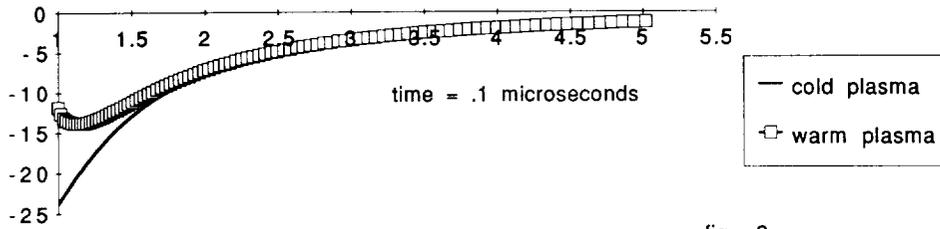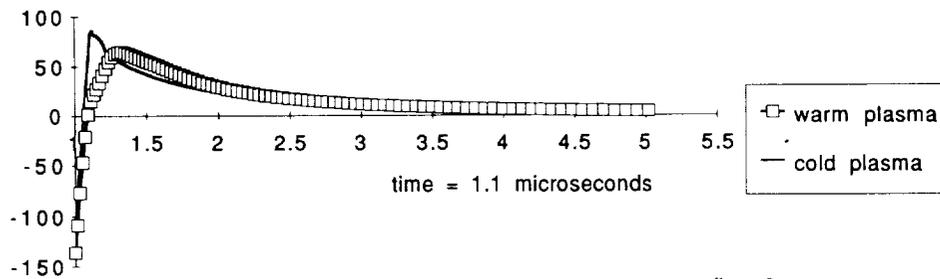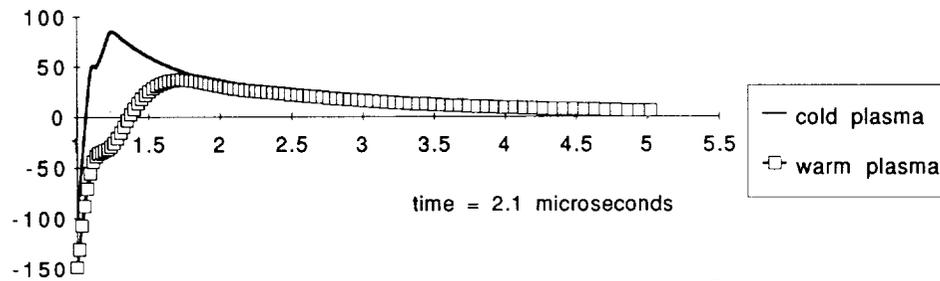
time = 3.1 microseconds
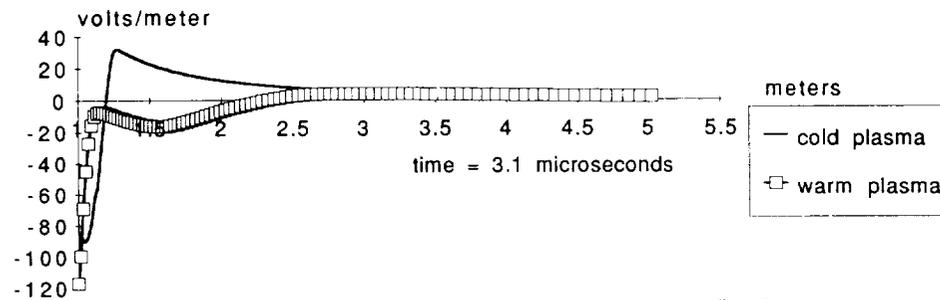
fig. 5

ELECTRON DISTRIBUTION FUNCTION AT 10th GRID POINT
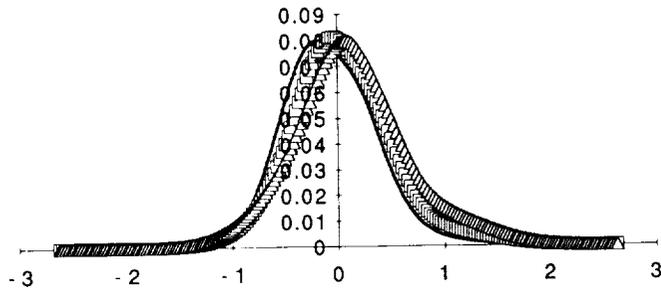
fig. 6

— 2.95	microseconds
-⊡- 3.00	microseconds
-△- 3.05	microseconds

velocity in 10**6 m/sec
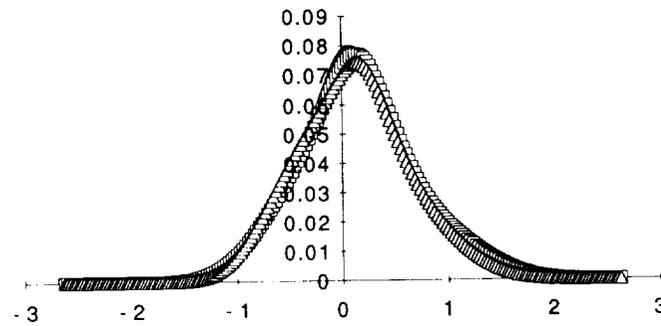
fig. 7

— 3.05	microseconds
-⊡- 3.10	microseconds
-△- 3.15	microseconds

velocity in 10**6 m/sec
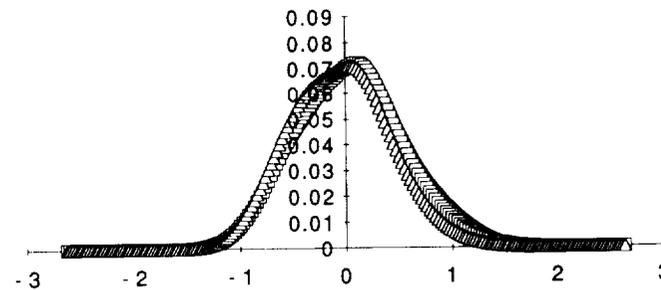
fig. 8

— 3.15	microseconds
-⊡- 3.2	microseconds
-△- 3.25	microseconds

velocity in 10 **6 m/sec
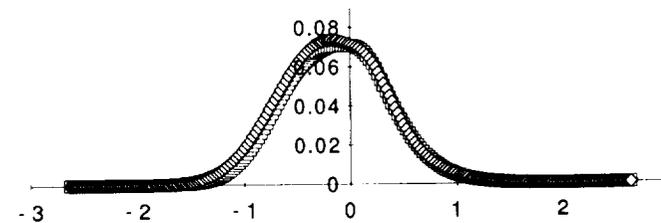
fig. 9

— 3.25	microseconds
-⊡- 3.3	microseconds
-◇- 3.35	microseconds

velocity in 10**6 m/sec

687

DISTRIBUTION FUNCTION BROADENING

fig. 10

I
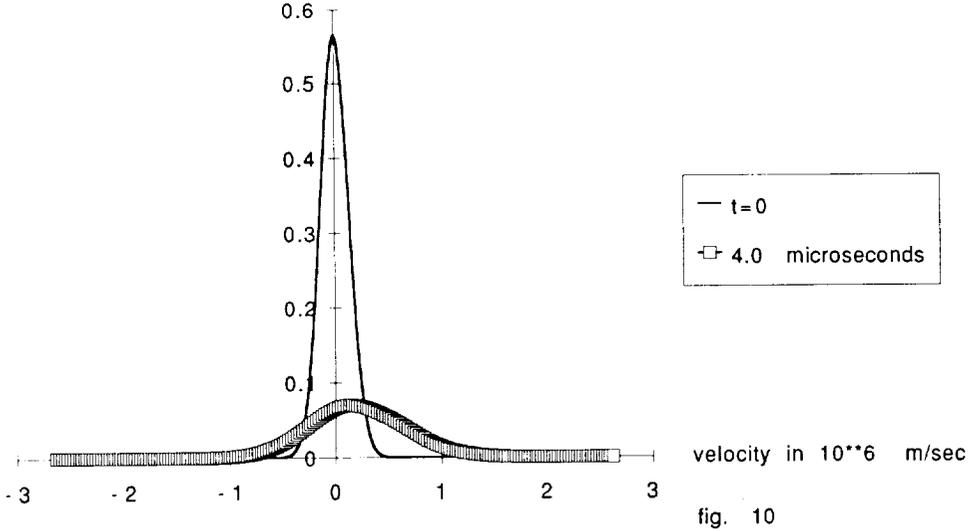
Erratum

The author has discovered that the warm plasma model discussed in this paper does not include angular momentum effects correctly. A corrected version has been constructed, and will be made available upon request.

Validation and Applications of the POLAR code

David L. Cooke,
Phillips Laboratory/ WSSI, Hanscom AFB, MA, 01731

Ira Katz, Gary Jongeward, John R. Lilley Jr.,
Scubed, LaJolla CA, 92038

Maurice Tautz, Radex Inc, Bedford MA, 01731

The POLAR code (POtentials of Large objects in the Auroral Region) was written to model the charging of large spacecraft by auroral electrons in low polar orbit. It has also proven to be a powerful tool for analysis of active and high voltage - plasma interactions. POLAR combines three dimensional object definition, direct solution to Poisson's equation, space charge transport, and surface charge computation. Physical effects treated by POLAR include the plasma wake and ram, the geomagnetic field, ionization, sheath and barrier formation, and photo and secondary electron generation and transport, POLAR development is complete and it is now available to the aerospace community. This talk will review the methodology of POLAR, but concentrate on the applications of POLAR to physical problems and observations with the intent of demonstrating both the code's capabilities, and its limitations. These applications include the Space Shuttle, the DMSP satellites, sounding rockets, plasma chamber experiments, and things not yet flown.

APPLICATION OF ENGINEERING CODES TO THE
ASSESSMENT OF SPACECRAFT CHARGING INDUCED HAZARDS

by
N. J. Stevens and M. R. Jones
TRW Space and Defense Sector
One Space Park
Redondo Beach, CA 90278

One of the major hurdles to developing a better under-standing of the environmentally-induced interactions in spacecraft systems (spacecraft charging) is the lack of simple tools to analyze the behavior of the spacecraft in a sub-storm environment. The available tool, "NASCAP", is expensive to use and treats only half of the problem - surface effects only. In response to this need, we have developed, under our IRAD program, a set of engineering tools that operate on personal computers and addresses both the surface and bulk charging aspects of the interactions. These engineering codes utilize circuit analysis techniques and model simplifications to allow rapid computations. The codes have been compared to both NASCAP predictions and to bulk charging data with excellent agreement for both phenomena.

The models used in our engineering codes are shown in Figure 1. NASCAP requires the use of a fixed grid and the definition of each surface in that grid. The computations then consider the particle trajectories from the grid point to space. Each grid point must be considered. In the engineering codes, the models are defined in terms of plates or triangles in an unconfined space. The predictions are based upon current balances to each plate. The code has routines to allow interactions between the plates to simulate "barrier" effects. The agreement for the predicted differential surface voltages for the model shown is given in figure 2. As shown, the agreement is excellent. Similar comparisons are available for the bulk charging predictions.

In conducting the comparison of engineering code to NASCAP predictions, it was found that some inconsistencies existed in the NASCAP code that are not obvious to the user. The use of the grid imposes constraints that voltage gradients exist between the grid points. This allows the existence of voltage gradients in conductors and influences surface resistivity predictions. These points will be discussed in the presentation along with the predictions for treating these inconsistencies properly. The model predictions for a variable intensity substorm will also be presented. This case is believed to be more applicable than a constant intensity substorm environment.

Session E3: INTERACTION LABORATORY EXPERIMENTS -I

Session Chair: R. Carruth

ION COLLECTION FROM A PLASMA BY A PINHOLE

David B. Snyder
NASA Lewis Research Center, Cleveland, OH

Joel L. Herr
Sverdrup Technology, Inc., Cleveland, OH

ABSTRACT

Ion focusing by a biased pinhole is studied numerically. Laplace's equation is solved in three dimensions for cylindrical symmetry on a constant grid to determine the potential field produced by a biased pinhole in a dielectric material. Focusing factors are studied for ions of uniform incident velocity with a three dimensional Maxwellian distribution superimposed. Ion currents to the pinhole are found by particle tracking.

The focusing factor of positive ions as a function of initial velocity, temperature, injection radius and hole size is reported. For a typical Space Station Freedom environment (i.e. oxygen ions having a 4.5 eV ram energy, 0.1 eV temperature and a -140 V biased pinhole), a focusing factor of 13.35 is found for a 1.5 mm radius pinhole.

INTRODUCTION

Present designs for Space Station Freedom SSF will result in structure potentials negative from the ambient plasma by about 150 V. This gives rise to concerns about sputtering by collection of ambient ions. On a broad scale this effect will be a contamination concern, but should not directly damage SSF structure. However SSF structure is covered by an insulating oxide layer. If holes develop in the insulation, ions will be focused into the hole and the local sputtering rate enhanced. At these pinhole sites sputtering will be aggravated. It is the enhanced collection of ions at these sites which this paper discusses.

Considerable effort has already gone into understanding how electrons are collected from plasma. Experimental measurements have been performed (ref. 1, 2), and computational models have been developed (ref. 3-5) to understand the 'snapover' effect, where electron collection to pinholes is significantly enhanced at potentials of a few hundred volts positive. However since ion collection does not exhibit this effect (due to the lower secondary electron yields) the simpler problem has not been investigated as intensively.

Recently, as part of the Space Station Electrical Grounding Tiger Team effort, Vaughn (ref. 6) has measured focus factors for ion collection, and Katz et al (ref. 7) have performed some initial calculations.

The model presented here does not solve Poisson's equation. Instead it solves Laplace's equation with computationally convenient boundary conditions. Thus it represents a simplification of the actual problem. While the results will not be exact, the model permits considerable qualitative examination of issues, and quantitative estimates of sputtering rates.

MODEL

The approach taken in this work is to solve Laplace's equation for a pinhole geometry, then track ion trajectories through the potential field. Electric fields are calculated from this potential field, and the equations of motion solved to find the trajectories of particles attracted toward the hole. The collection enhancement (focusing factor) can be found by finding the radius where particles no longer hit inside the hole. This approach does not include space charge effects. An approximation for the plasma sheath is included by adjusting the height of the calculation space.

Potential Fields

The potential fields were found by using the Point Gauss-Siedel method to solve the cylindrical difference formulation of Laplace's equation in a rectangular grid. The depth of the hole is not included in the model. Instead it is considered to be a uniformly biased conductor. While this may not be reasonable for modeling debris damage, it is reasonable for modeling the effect of insulation damage when the insulation thickness is small compared to the hole radius. The

.PRECEDING PAGE BLANK NOT FILMED

insulation is treated as being at the plasma potential, 0 V. This results in a lower boundary having two potentials. Inside the hole radius the potential is at the bias potential. Outside, it is zero, referenced to the plasma potential. The upper boundary is set to the plasma potential, zero.

The potential contours in figure 1 show that the potential drops off very rapidly, and the potential contours that extend large distances are of low magnitude. The potential at about 1 radius from the center of the hole drops to 31 % of the applied voltage. The potential varies most rapidly at the edge of the hole. This is the region of strongest electric field. Near the hole edge, for a negative potential hole, the electric field flux lines go up from the surface and down to the hole surface. This shows up in the trajectories of particles which nearly miss the hole edge but are forced up and into the midregion of the hole.

In order to satisfy Dirichlet-Neumann conditions and completely specify the problem, the inner ($r=0$) and outer boundaries must also be specified. The inner boundary is a center of cylindrical symmetry, so the radial component of the electric field, E_r, must be zero. The electric field on the outer boundary also approaches zero as the radius approaches infinity. For these calculations E_r is also set to zero at this boundary. Unfortunately, the trajectory of particles near the boundary are very sensitive to the boundary's existence. Particles near the outer boundary tend to fall straight to the surface. However, when the boundary is moved away those particles may have a considerably different trajectory. For this work the boundary was set far enough away from the hole that the focusing factor did not change. The trajectory of particles not hitting the hole will be incorrect, but that aspect of the problem will be irrelevant to the results of this calculation.

Electric Field

The electric fields obtained from the potential grid will be integrated in the equations of motion to yield particle trajectories. Therefore the interpolated electric field within each grid region must be consistent with a relatively continuous potential field. Discontinuities in the potential field between grid regions can cause the integrated kinetic energy accumulated by the particle to disagree with that expected from the potential at a point and energy will not be conserved. This turns out to be especially important near the edge of the hole where the potential changes rapidly in one grid spacing and the electric field is both large and rapidly varying with position.

To satisfy this criterion, the electric field is evaluated in the center of the grid, and using additional interpolated potential points the gradient of the electric field components are evaluated as illustrated in figure 2. Interpolated potential points are evaluated to satisfy Laplace's difference equation. That is, they are the average of the four surrounding points. Potentials at the

center of the grid regions, $P(r,h)$ are obtained from the potential grid, $P(j,i)$. The potential at the center of the grid edge, $P(r,i)$ can then be interpolated from the two nearby grid centers and the two nearby grid corners.

The electric field components at the center of the grid can now be evaluated from the grid edge potentials and the gradients can be evaluated using the corner potentials. The electric field can then be interpolated anywhere in the grid region. The electric fields calculated in this way are pinned to reproduce the corner potentials as well as the center of the grid edge. The resulting potential field is continuous between grid regions through these points, and constrains the rest of the edges from being too discontinuous.

This method does not add new information to the known values at the grid points. Rather it uses the known values, and interpolates between them in a relatively continuous manner. The procedure could be reiterated to generate as continuous a potential field as required.

Particle Tracking

Particle Trajectories are found by integrating the equation of motion, $a = F/m$ to obtain velocity (v) and position (r). The electric field, E, is evaluated either for calculation purposes relative to the center of the grid region, r_0, or, for integration purposes relative to the position at the beginning of a time step, $r(0)$,

$$E(r) = E(r(0)) + (r-r(0)) \bullet \nabla E.$$

The acceleration is given by,

$$a = q/m \, [E(r(0)) + (r-r(0)) \bullet \nabla E]$$

the velocity is obtained from,

$$v(T) = v(0) +_0\!\int^T a \, dt.$$

The first order expansion of E may be used to replace a. An integrable estimate of r can be obtained by noting that,

$$r(t) \sim r(0) + v(0)t + q/2m \, E(r(0))t^2,$$

as long as the term $(r(t)-r(0)) \bullet \nabla E$ is small compared to $E(r(0))$.

v can be iterated using,

$$v(T) = v(0) + q/m \, E(r(0))T$$

$$+ \,_0\!\int^T [v(0)t + q/2m \, E(r(0))t^2] \bullet \nabla E \, dt$$

finally,

$$v(T) = v(0) + q/m \, E(r(0))T$$

$$+ q/m \, [v(0)T^2/2$$

$$+ q/6m \, E(r(0))T^3] \bullet \nabla E$$

similarly,

$$r(T) = r(0) + v(0)T$$

$$+ q/2m\ E(r(0))T^2$$

$$+ q/6m\ [v(0)T^3$$

$$+ q/4m\ E(r(0))T^4] \bullet \nabla E$$

It is possible to take fairly large spatial steps using this formulation. The main constraint is to prevent the particle from moving too far into the next grid region where the extrapolation is not valid. In principle the particle could cross the grid in one large time step, then be bumped across the edge in the next step. In practice, particles were moved about 0.2 grids each time step. An estimate of the time step size can be made by solving for T in the quadratic equation,

$$r(T) = r(0) + v(0)T$$

$$+ q/2m\ E(r(0))T^2,$$

where $r(T)$ is evaluated to cross the radially symmetric grid in all six directions and the smallest positive T is actually used. Particles which pass through the $r = 0$ line are reflected by changing the sign of the radial component of the velocity.

The potential field around the hole has cylindrical symmetry, so the angular coordinate becomes important only when visualizing the actual three dimensional trajectory. However, the angular velocity component plays a significant role. Angular effects were accounted for by calculating the motion in three dimensions, then rotating the coordinate system to find the new r, z, v_r, v_z, v_θ. This technique was tested by tracking a particle with no electric field present and ensuring that it followed a straight line.

Implementation of model

An objective of this effort was to produce a fast microcomputer model which could be used to investigate ion collection by pin holes. Figure 3 shows the flow of the program used.

Solving the potential field takes a significant amount of time, 10 minutes on an Intel 80387 to solve an 80 X 160 grid, but this need only be done once. The time can be minimized by using as limited a computational grid as possible. Solution of the particle trajectories is fast due to the semi-analytical approach taken. Typically times steps were scaled to cross 0.2ths of a grid region.

The approximation for r used above to enable integration of the acceleration begins to break down when $(r(t)-r(0)) \bullet \nabla E$ is not small compared to $E(r(0))$, as is the case near the hole edge. However when this occurs the particle is almost certainly collected, though the calculated trajectory and impact point may not be accurate. A calculation space with a radial dimension

twenty times the hole size gives sufficiently accurate results.

Velocity Distribution

The effect of a Maxwellian velocity distribution can be addressed by adding incremental velocities to the incident velocity representative of the three dimensional velocity distribution. The results for each velocity element are weighted by the probability of a particle having an incident velocity near the tested velocity.

To calculate a focus factor for a given temperature, the three dimensional particle velocity space was broken into a grid. Division of the velocity space into approximately 3000 (18x18x9) elements yielded a reasonable simulation. Only half the θ velocity is needed due to the symmetry of the problem. Each grid element was assigned a velocity vector representative of its position and a weight of the fraction of particles contained in it. Focus factors for each velocity grid element were then obtained. The net focus factor is then the sum of the focus factor/weight product.

DISCUSSION

Focus Factor

For particles whose initial velocity is normal to the surface, the focus factor is easy to find. All particles emitted from the upper boundary inside a certain radius, r_c, will be collected. The focus factor, that is the enhancement in total current collected is the ratio of the collecting area to the hole area, or $(r_c/r_h)^2$.

It is interesting that the impact points do not map monotonically with the incident radius. As shown in figure 4, both for particles near the center and far from the edge the further away from the center a particle is dropped, the further away it hits. But particles which traverse near the edge can end up almost anywhere.

The following dependencies were investigated for a base set of conditions simulating RAM impact conditions for SSF orbits, i.e.: oxygen ions, at normal incidence, with an initial velocity corresponding to 4.5 eV, and with a temperature of 0.1 eV.

Energy Dependence

A qualitative argument can be made to describe the focus factor's dependence on incident energy. In this model both the upper boundary (plasma sheath edge) and the lower boundary (surface) have the same potential. Therefore if the incident particle picks up more kinetic energy, due to motion directed parallel to the surface, than its initial incident energy, it will not be able to reach the lower surface except in the hole. If a particle is dropped into the grid with zero velocity and

696

temperature, except at the outside edge, it will reach the hole and the focus factor is essentially infinite.

In low earth orbit the focus factors will be smaller than those which are observed in ground tests since the ram energy of the incident ions is larger than the thermal energy of ions used in most ground tests.

Results of the calculated focus factor as a function of incident velocity are shown in figure 5.

Hole Size and the Plasma Sheath

The calculations presented here scale with two geometric parameters, the hole size relative to both the width and height of the calculation space. To obtain meaningful results the hole size should be small compared to the total surface area. Therefore the calculation grid should be wide enough that further expansion of the grid does not affect the trajectories of those particles which hit or nearly hit the hole.

On the other hand, the height of the calculation space indicates a distance above which the electric field is zero. Therefore it performs a function similar to the plasma sheath. These results obtained by changing the pinhole radius as compared to the sheath thickness are displayed in figure 6.

It is expected that as the hole size becomes large compared to the sheath the focus factor will approach unity. But in this case the height of the calculation space would be identified as the Child-Langmuir length. As the hole gets small compared to the height of the calculation space the focus factor will approach a value which depends only on incident angle and temperature, and incident energy. However, in these simulations, as the hole gets even smaller the focus factor drops off again. Possibly the particles cannot reach the hole due to angular momentum constraints.

Temperature

As indicated above, modeling temperature effects requires finding focus factors for a large number of initial velocities. Particles with high angles of incidence may hit the hole from considerable distances away. Fortunately these represent a small portion of the velocity space. The effect of temperature is shown in Figure 7 for two cases. Figure 7a shows a case where the temperature is small compared to the incident velocity. Figure 7b shows a case with zero incident velocity.

Comparison to Experiment

Vaughn (ref. 6) has measured ion currents to a hole. He measured the current to a 0.54 cm diameter hole biased at -140 V. The Argon plasma he used had a number density of 2×10^{12} m^{-3}, an electron temperature of

1.2 eV and an ion energy of 2.0 eV. An ion current of 0.5 μA was measured. This suggests a focus factor of 78, if an ion thermal current is used, to 22, if the ions are assumed to be flowing. Since there was a 2 V drop from the plasma source to the tank plasma, we assume the latter number is the correct value.

The sheath thickness for comparison was estimated by finding where $\partial E_h / \partial h$, from an average potential seen as a function of height, is equal to ϱ / ε_o from the plasma density. A sheath thickness of .018 m was estimated for the ram case, and .021 m for the thermal case. This model predicts a focus factor of 18 for the ram velocity case, and 8.5 for the thermal case. The agreement is good.

Katz et al (ref.7) have noted that an accurate model must include a good model of the sheath edge, and also must model the surface potentials near the hole accurately. They have noted that the surface near the hole is shielded from thermal electrons by the hole's negative potential and tends to charge positive.

CONCLUSION

The simple pinhole ion collection model presented here permits reasonable estimates of focus factors. The model uses a solution to Laplace's equation of the cylindrically symmetric potential field to evaluate electric fields near the hole, then tracks particles to evaluate the focus factor. The particles are tracked in three dimensions, though the angular coordinate can be ignored. The model permits incident velocities in three dimensions and can therefore be used to model temperature effects.

An objective of this work was to provide a fast model pin hole collection. Focus factors for specific initial velocities are found quickly but temperature calculations are time consuming due to the large number of particles tracked.

The present model assumes a level plasma sheath and a uniform surface potential. The accuracy of the model might be improved by including an analytical model of the sheath shape and the surface potential. We intend to verify the accuracy of the model further by comparing its result with I-V curves obtained in the ion collection tests.

REFERENCES

1. Stevens, N.J., Berkopec, F.D., Purvis, C.K., Grier, N., and Staskus, J., "Investigation of high voltage spacecraft interactions with plasma environments," AIAA/DGLR 13th International Electric Propulsion Conference Paper 78-672 (1978).

2. Gabriel, S.B., Garner, C.E. and Kitamura, S., "Experimental Measurements of the Plasma Sheath Around Pinhole Defects in a Simulated High-Voltage Solar Array," AIAA Paper 83-0311 (1983).

3. Chaky, R.C., Nonnast, J.H., and Enoch,J. "Numerical simulation of sheath structure and current-voltage characteristics of a conductor-dielectric in a plasma," J. Appl. Phys. 52(12), pp. 7092-7098 (1981).

4. Brandon, S.T., Rusk, R.L., Armstrong, T.P., and Enoch, J., "Self-Consistent Simulation of Plasma Interactions with Secondary-Emitting Insulators," Spacecraft Environmental Interactions Technology 1983, NASA CP 2359, pp 287-303 (1985).

5. Mandell, M.J., and Katz, I., "Surface Interactions and High-Voltage Current Collection," Spacecraft Environmental Interactions Technology 1983, NASA CP 2359, pp 305-319 (1985).

6. Vaughn, J., "Ion Focusing Test", Electrical Grounding Tiger Team Meeting, Huntsville, Al, May 14-17, 1991.

7. Katz, I., Davis, V.A., and Mandell, M.J., "Ion Focussing and Implications for SiO2 Deposition in Solar Array Gaps," Electrical Grounding Tiger Team Meeting, Huntsville, Al, May 14-17, 1991.

potential contours for a -140 V bias
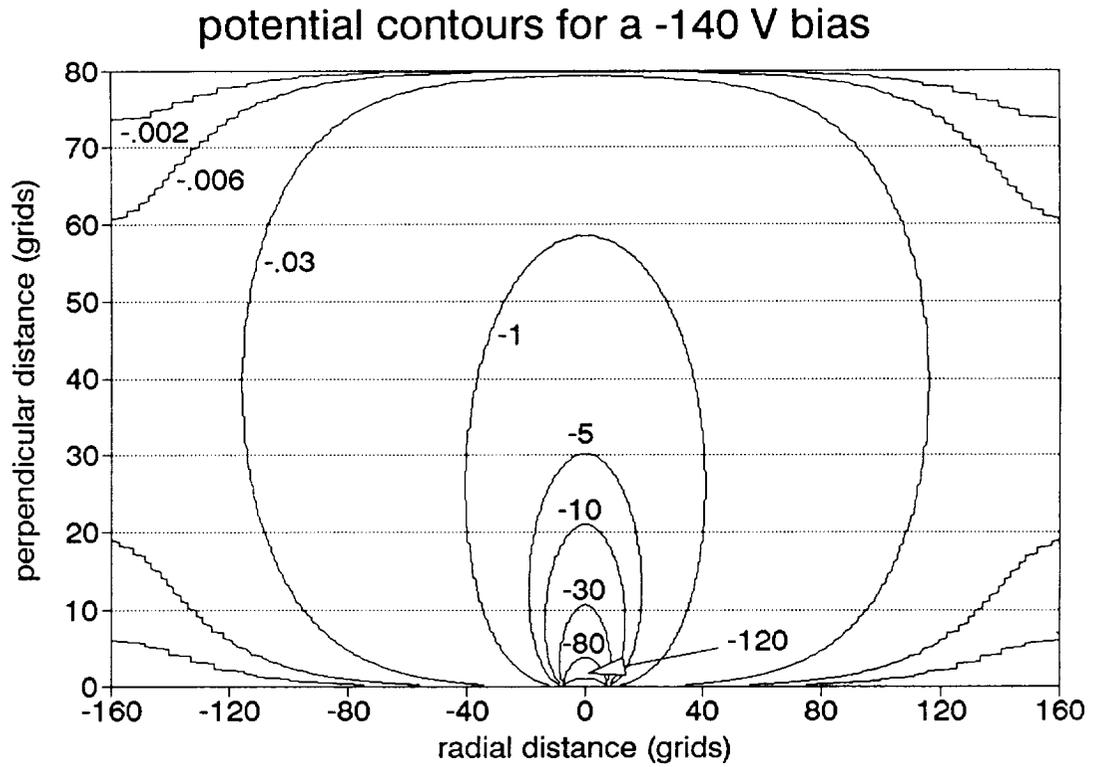
Figure 1. Potential contours near hole. Hole bias is -140 V and hole radius is 8 grids.
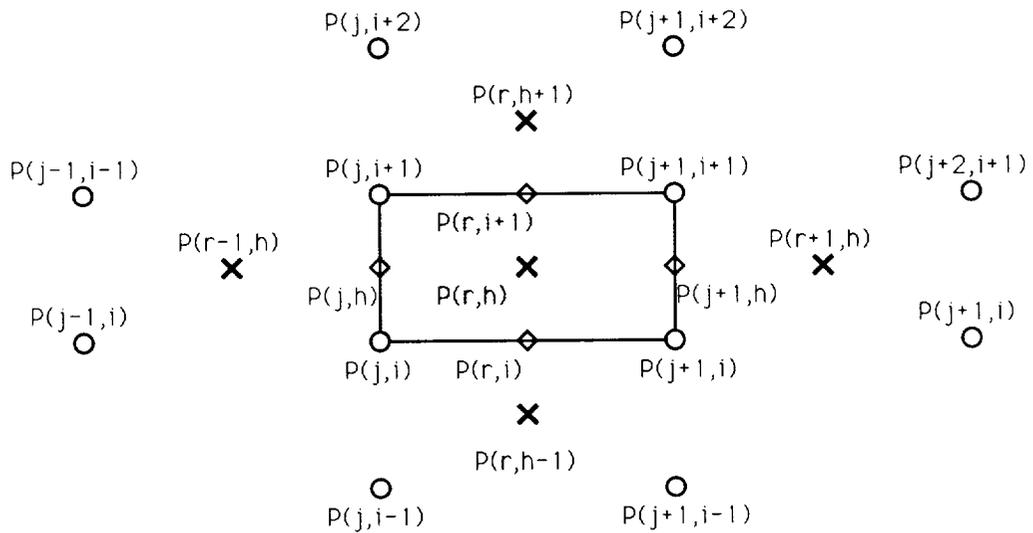
Figure 2. Potential interpolation scheme.
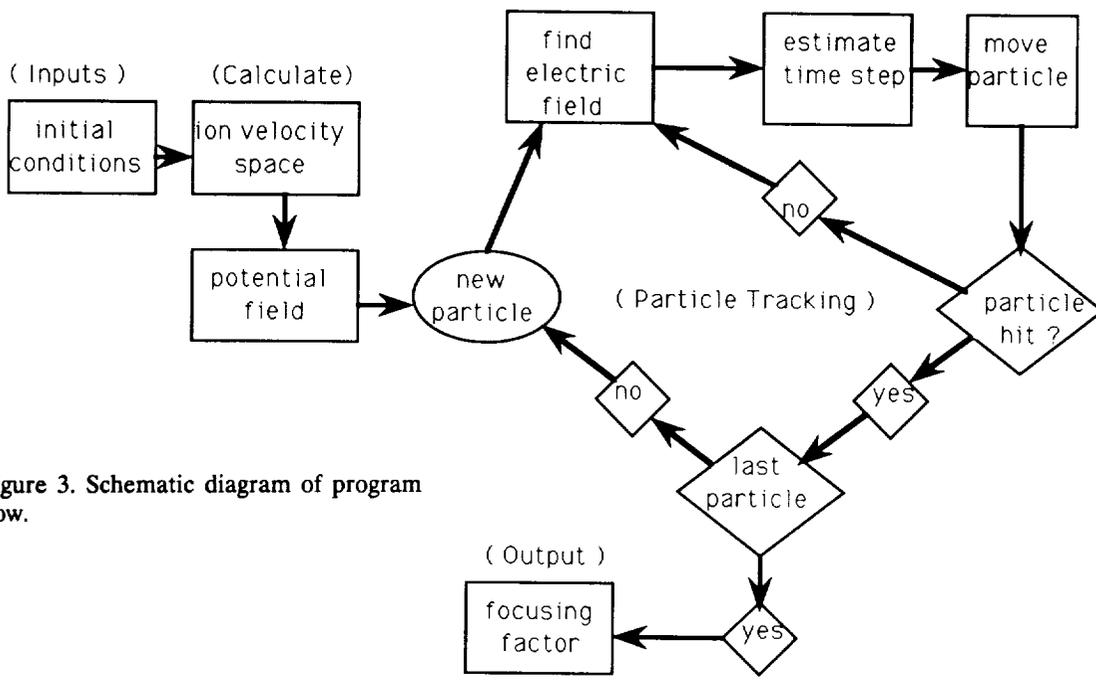
Figure 3. Schematic diagram of program flow.

impact points versus incident radius
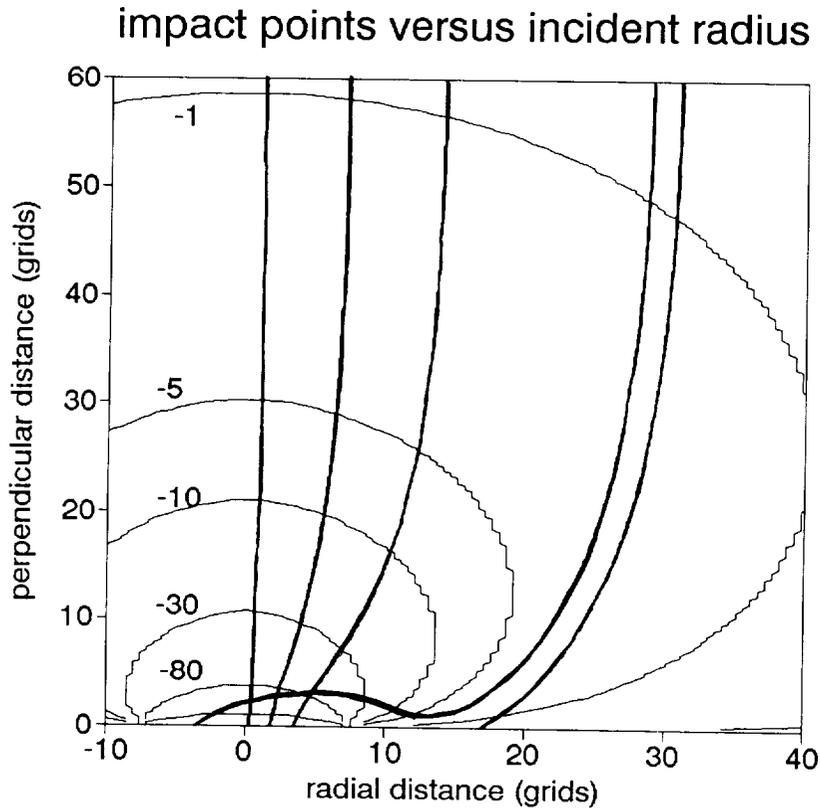
Figure 4. Examples of particle trajectories.

incident energy variation
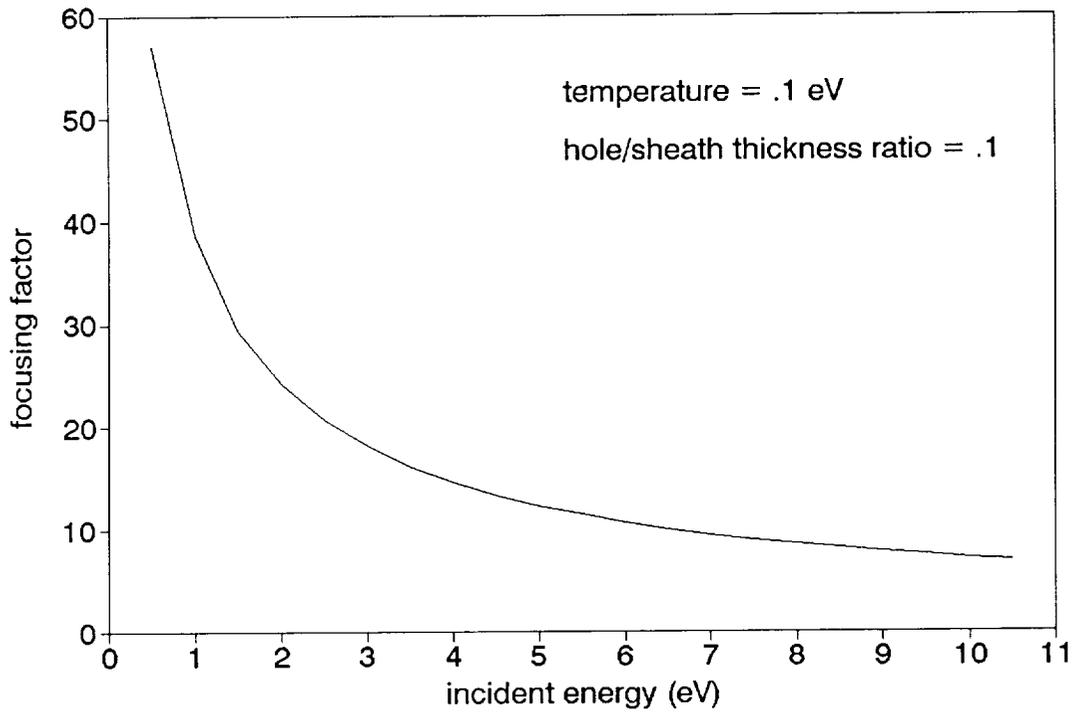
Figure 5. Relationship between energy of incident particles and resulting focus factor.

pinhole radius variation
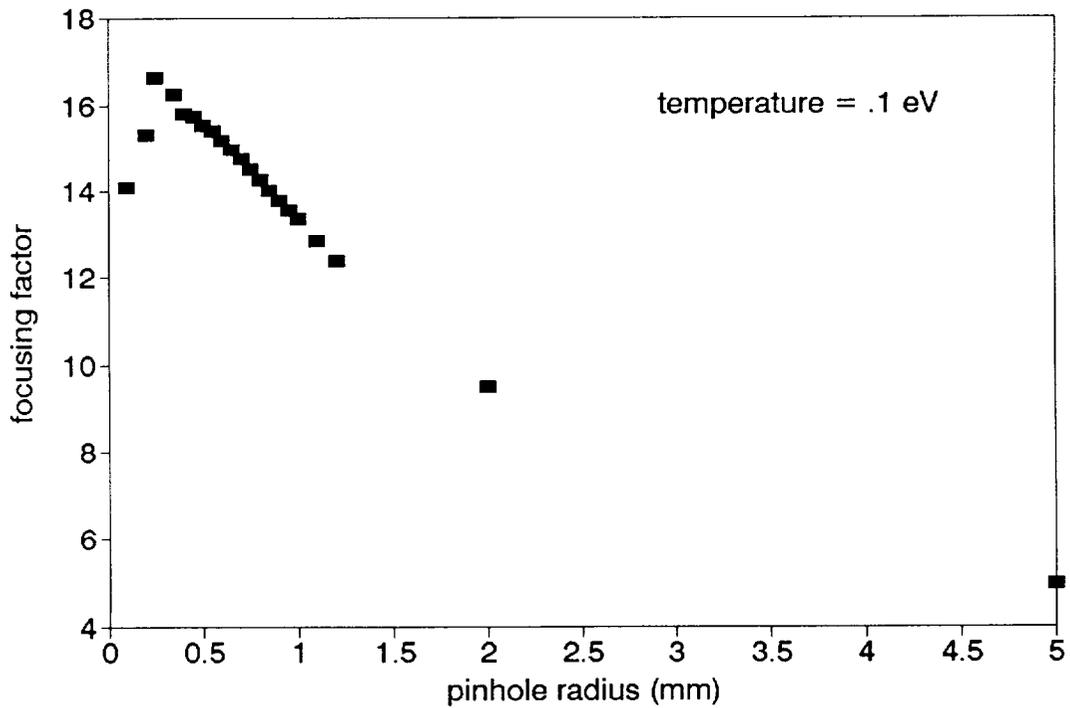
Figure 6. Effect of pinhole size on focus factor.

701

Maxwellian temperature variation

Figure 7a. Effect of temperature on focus factor.

Maxwellian temperature variation
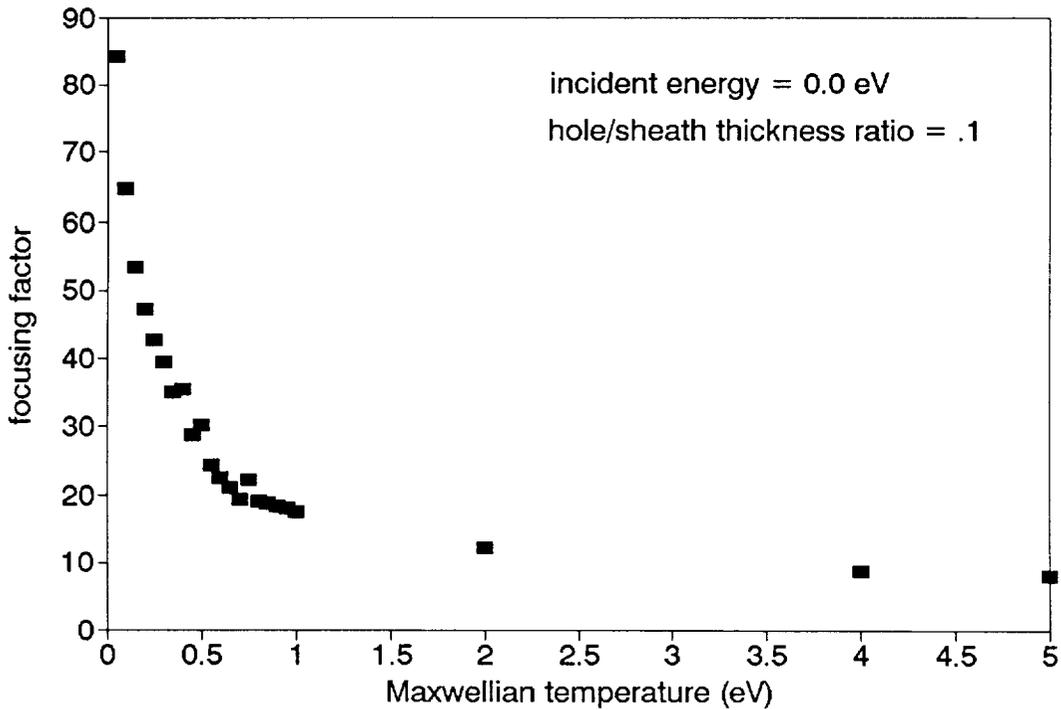
Figure 7b. Effect of temperature on focus factor.

702

EXPERIMENTAL BREAKDOWN OF SELECTED ANODIZED ALUMINUM SAMPLES IN DILUTE PLASMAS

Norman T. Grier and Stanley Domitz
NASA Lewis Research Center
21000 Brookpark Road
Cleveland, Ohio 44135

ABSTRACT

Anodized aluminum samples representative
of Space Station structural material were
tested for electrical breakdown under space
plasma conditions. In space, this potential
arises across the insulating anodized
coating when the spacecraft structure is
driven to a negative bias relative to the
external plasma potential due to plasma-
surface interaction phenomena. For
anodized materials used in the tests
reported herein, it was found that break-
down voltage varied from 100 volts to 2000
volts depending on the sample. The current
in the arcs depended on the sample, the
capacitor, and the voltage. The level of the
arc currents varied from 60 to 1000
amperes. The plasma number density
varied from 3×10^6 to 10^3 ions per cc. The
time between arcs increased as the number
density was lowered. Corona testing of the
anodized samples revealed that samples
with higher corona inception voltage had
higher arcing inception voltages. From this
it is concluded that corona testing may
provide a method of screening the samples.

INTRODUCTION

Anodized aluminum surfaces are proposed
for use as the outer skin for many of the
components on Space Station Freedom
(SSF). These structural components are
electrically connected to the negative
terminal of the SSF solar array. The
proposed operating voltage for the solar
array is 160 volts. Due to the plasma-array
interactions, the spacecraft structure is
predicted from numerical simulations
calculations to be driven to approximately
140 volts negatively with respect to the
surrounding plasma potential. Since
anodized aluminum is an insulator, the
plasma-anodized-aluminum form a
capacitor with the plasma and the
aluminum forming the two electrodes.
Thus, the anodized layer is expected to
withstand a voltage of 140 volts during the
spacecraft operations. Breakdowns of this
thin anodized layer can initiate arcs with
the current being supplied by the discharge
of the large plasma-anodized-aluminum
capacitor. The interaction of charged
spacecraft surfaces with the space
environment at low earth orbit has been the
subject of many studies (references 1-8).
The studies, however, are primarily focus
on determining the current to the surface
from plasma rather than the breakdown of
insulators in plasmas.

Previous experiments at Marshall Space
Flight Center (MSFC) have demonstrated
that some anodized coatings breakdown
and arc at voltages as low as 80 volts in
plasma, causing very large currents in the
circuit. (References 9 and 10.) This result
was unexpected since "normal" arcs which
occur at metal-insulator interfaces, are
initiated at voltages greater than 200 volts
and usually in the range of 300 to 1000
volts. The present experiments were
performed to extend the measurements
obtained at MSFC to other anodized
aluminum samples. Samples tested here

included two samples that were obtained from MSFC along with other samples. All the other samples had breakdown strengths greater than 200 volts. The two MSFC samples had breakdown strengths less than 200 volts. These results agree the those obtained at Marshall for these type of samples. The thermal and optical properties of the samples were not measured, therefore, the relative merits of one over the other for SSF thermal control purposes was not discerned.

There apparently are two type of arc sites, (1) virgin sites that are initiated during the breakdown process, and (2) repeating sites that are initiated at previous breakdown sites. Each type seem to have its own threshold voltage. It was expected that if arcs are initiated at previously arc sites, the threshold voltages at these sites would increase as the plasma number density is lowered. For the JSC and the SSF samples, this was found to be true. It is therefore concluded that the arcs occurring on these samples, mostly occur at previous breakdown sites. For the MSFC samples the arc inception voltage did not vary with density. It is concluded from this that each arc produced a new breakdown site. And due to the longer time for the anodized surfaces to become charged to the applied voltage, a lower density cause the arc rates to be reduced . The measurements were performed at number densities ranging from 3×10^3 to 3×10^6 e/cc.

It would be advantageous to perform bench tests of anodized samples that are predictive of their behavior in a plasma. Corona (partial discharge) testing is a nondestructive test that can usually predict insulator breakdown behavior when ordinary metallic electrodes are used. The corona inception voltage is always lower than breakdown voltage. Corona tests were performed on each of the samples used herein. A correlation between corona inception voltage and breakdown voltage was found.

EXPERIMENTAL TESTS

The tests were performed in the NASA LeRC vertical vacuum chamber, which was oil diffusion pumped to a base pressure of 2×10^{-7} torr. The plasma sources were an argon hollow cathode that are commonly used in ion thrusters and/or and two Penning type hot filament discharge chambers. When both sources were operated simultaneously, the plasma densities ranged from 1 to 3×10^{-6} e/cc. To obtain lower densities, the hot filament Penning type chambers were used alone. These sources are controllable down to very low densities, and provided number densities from approximately 1000 to 3×10^5 electrons per cubic centimeter. The plasma density was measured using spherical Langmuir probes. The anodized aluminum samples were attached to an approximately 12-inch long Kapton coated wire pigtail which were in turn connected to the center conductor of a coaxial cables. A sketch of the experimental setup is shown in Figure 1. All the exposed edges and connections to the samples were covered to prevent break-down and current collection in these areas. To test the integrity of the seal, the samples were biased from -200 to either +200 or +300 volts in a plasma and the current measured.
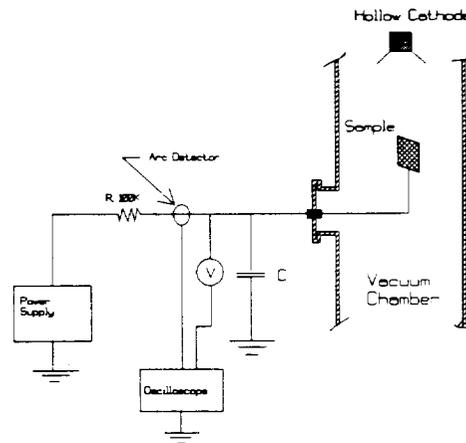
Figure 1 Test Setup

Table I Description of Samples Tested

Sample	Size	Electrolyte	Thickness, mils
MSFC-1	6x6 in	Chromic	0.20
MSFC-F	6x6 in	Chromic	0.20
JSC-1	3x3 in	Sulfuric	0.99
JSC-2	2x2 in	Sulfuric	0.20
JSC-3	6x6 in	Sulfuric	0.20
SSF-M1	3x3 in	Sulfuric	0.20

A typical test was initiated by introducing the argon plasma and applying a negative bias to the anodized sample until an arc occurred. The arc current was detected using a Pearson current transformer. The output was fed to one channel of a Hewlett Packard oscilloscope. The other channel of the oscilloscope received the signal from a Textronic 100:1 voltage probe, which was used to measured the voltage applied to the sample. The capacitors used in the circuit ranged from .25 µf to 100 µf. They were high voltage dry film type. All tests were performed at a plasma density of 2.5×10^{6} e/cc. To determine the effect of density on the arc inception voltage, additional tests were repeated using a capacitor of 25 µf at lower plasma densities.

Corona measurements were also made on each sample using a Biddle Partial Discharge Detector with a sensitivity of 1 picocoulomb. The corona measurements were made using a flat metal electrode to provide the electrical contact to the anodized layer. Realizing that plasma electrodes may perform differently than metal electrodes, attempts were made to perform the measurements using an electrolyte electrode, in this case salt water, as one of the electrodes. It was hoped that this would simulate the plasma as being one the electrodes. In some tests, these measurements were not useful since at times they indicated a corona inception voltages less than 20 volts less than the sensitivity of the instrument.

TEST SAMPLES

The test samples are listed in Table I. The two samples obtained from MSEC were used as a baseline for comparing test results. The SSF sample is a sample of one of anodized aluminum that is being considered for use by LeRC on SSF. Samples JSC-1, -2, and -3 were obtained from NASA Johnson Spaceflight Center (JSC) and were made from aluminum 5657 and used sulfuric acid as the electrolyte. These samples were pretreated for 20 seconds in bright-dip at 97°C, anodized in 20% H_2SO_4 and sealed in water at 95°. The two SSF samples were made of Type IIClass 1 aluminum 6061T6 and sulfuric acid was also used as the electrolyte. The anodization was done on these samples according to Mil Spec 8625-E.

TEST RESULTS

It was found that breakdown voltage of the anodized aluminum samples varied considerably in a plasma, from 100 volts to over 2000 volts. The maximum current varied from 60 to 1100 amperes. The current depended on the sample, the capacitance, and the applied voltage. Table II summarizes the tests the results. Figure 2 shows a typical oscilloscope trace obtained during an arc.
Results were complicated by the fact that two types of arc discharges can occur. One initiated by the breakdown of the anodized film and the other, the "normal"

Table II Test Results

Sample	Inception, V	IMAX, A	I*dt, Coul	C, μF	CV
MSFC-1	200	5	1.81×10^{-4}	5.0	1.0×10^{-3}
(6x6 in)	200	20	4.13×10^{-3}	20	4×10^{-3}
	120	25	2.74×10^{-3}	25	3×10^{-3}
	180	100	1.10×10^{-2}	100	1.8×10^{-2}
MSFC-F	200	110	2.2×10^{-3}	20	4×10^{-3}
(6x6 in)					
JSC-1	600	80	6.0×10^{-4}	0.25	1.5×10^{-4}
(3x3 in)	600	100	1.0×10^{-3}	0.5	3.0×10^{-4}
	500	350	5.0×10^{-3}	5	2.5×10^{-3}
	800	600	9.0×10^{-3}	10	8.0×10^{-3}
	600	500	8.9×10^{-3}	25	1.5×10^{-2}
	700	1100	7.2×10^{-2}	100	7.0×10^{-2}
JSC-2	400	48	4.6×10^{-4}	0.5	2.0×10^{-4}
(2x2 in)	400	85	9.2×10^{-4}	1	4.0×10^{-4}
	400	160	2.9×10^{-3}	5	2.0×10^{-3}
	400	206	6.4×10^{-3}	15	6.0×10^{-3}
	400	293	1.8×10^{-2}	45	1.8×10^{-2}
	400	400	4.4×10^{-2}	100	4.0×10^{-2}
JSC-3	300	38	3.5×10^{-4}	0.5	1.5×10^{-4}
(6x6 in)	300	57	6.0×10^{-4}	1.0	3.0×10^{-4}
	300	120	2.1×10^{-3}	5	1.5×10^{-3}
	300	183	6.8×10^{-3}	15	4.5×10^{-3}
	300	320	1.7×10^{-3}	45	1.35×10^{-2}
	300	400	3.3×10^{-2}	100	3.0×10^{-2}
SSF-M1	600	48	5.9×10^{-4}	0.5	3.0×10^{-4}
(3x3 in)	600	146	1.4×10^{-3}	1.0	6.0×10^{-4}
	600	233	4.4×10^{-3}	5	3.0×10^{-3}
	600	360	1.3×10^{-2}	15	9.0×10^{-3}
	600	533	2.7×10^{-2}	45	2.7×10^{-2}
	600	626	6.5×10-2	100	6.0×10^{-2}

type occurring at bare metal sites that have become exposed after breakdown. This latter type occurs at the junction of a metal-insulator interface when it is biased negatively relative to plasma potential. the breakdown voltage of these two types can be quite different. For example, for the JSC-1 sample the initial breakdown voltage was approximately 2000 volts. However, subsequent arcing occurred at approximately 600 volts at the now exposed metal site. As a result, the potential required for film breakdown could no longer be reached and only one hole appeared in the insulator.

On the other hand, for MSFC-1 and MSFC-2 samples, the film breakdown strengths were in the 100 to 200 voltage range. These voltages were not high enough for breakdowns to reoccur at ametal site. Each arc, created therefore, a new site. For other samples both types of breakdowns were in the same range, so

Range   16 V   Auto Scale [Enabled]
Offset   2.000 V   Store Mode [Normal   ]
Probe   [1:1   ]   Coupling   [ac] [1 MΩ]

Graph [1]   2.00 V/div   2.00 V   50.0 μs/div   -180.0 μs

1: [Chan 1]
    I 100A/V

2: [Chan 2]
    V 400V/DIV
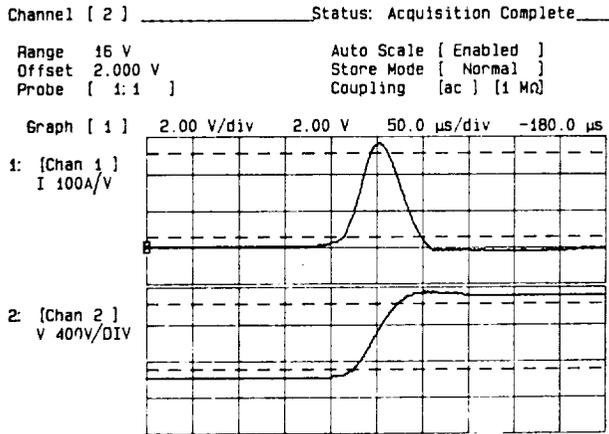
Figure 2. Typical oscilloscope traces.

Figure 3. The logarithm of the maximum current as a function of log of capacitance. C
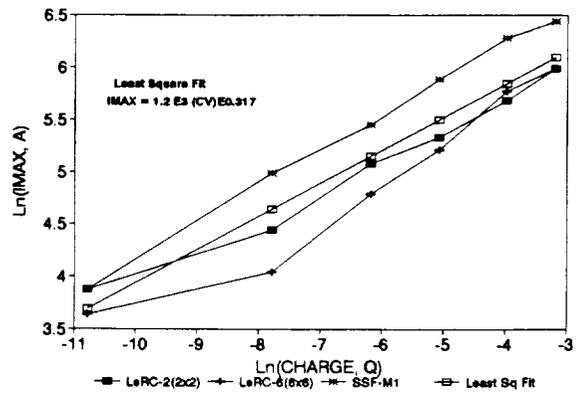
Figure 4. The logarithm of the maximum Current as a function of log of total charge. CV.

arcing occurred through both insulation breakdown and at bare metal sites. The maximum arc current depended on the capacitor size and the voltage reached at breakdown. Figures 3 and 4 show respectively the maximum current, I_{MAX}, as a function of the capacitance and the maximum current as a function of the total charge stored in the capacitor at discharge, CV. Both plots are log-log plots. The results indicate the maximum current varies as 0.317 power of either the capacitance, C, or the charge , CV.

From the transient current-voltage oscilloscope traces, (a typical one is shown in figure 2), it was found that the total integrated current is about the same as the charge stored on the capacitor before discharging. This implies that the capacitor was fully discharged in each arc. This is shown in Table II in columns labeled I*dt and CV. As can be seen, the larger the capacitor the better the agreement.

Comparing the maximum current measured from the oscilloscope traces and that found by using the maximum dv/dt showed relatively good agreement. The difference is attributed to the inductance in the circuit. The circuit inductance was determined to be in the range of 20-30 microhenries. It is felt that most of this inductance is due to the inductance in the capacitors.

Density Effects

By using the hot filament Penning type discharge sources, the density could be varied down to approximately $3x10^3$ e/cc. Tests were run to determine breakdown strength as a function of density. The results are shown in Table III. For the three samples shown, the breakdown strength increases with decreasing number densities. This indicates that arcing is occurring at bare metal sites. This agrees with previous investigations performed using negatively biased pinholes and solar arrays. (References 11 and 12.)

On the other hand arcs initiated by breakdown of the anodizing layer should be independent of the number density. The only requirement is that the thermal flux of ions to the surface be sufficient to keep the surface charged while also providing for leakage current through the

Table III Effect of Density on Breakdown Voltages

Density, e/cc	LeRC-3 V_{bk}	LeRC-2 V_{bk}	SSF-M1 V_{bk}
$2x10^6$	300	400	600
$5x10^4$	350	600	700
$7x10^3$	400	800	900

insulator and the opened arc sites. This is demonstrated for sample MSFC-F in Figure 5. The data for this figure was obtained by holding the voltage at -200 volts while the number density was lowered and recording the time between arcs at several density levels. the lowest density where arcs were observed was approximately $2x10^3$ e/cc. As can be seen, as the density increases arc rate increases sharply.

Sample: MSFC-F

Applied Voltage, - 200 volts

Capacitance, C, 25 microfarads

Time Between Arcs, Sec

Plasma Electrons Density, e/cc (Thousands)
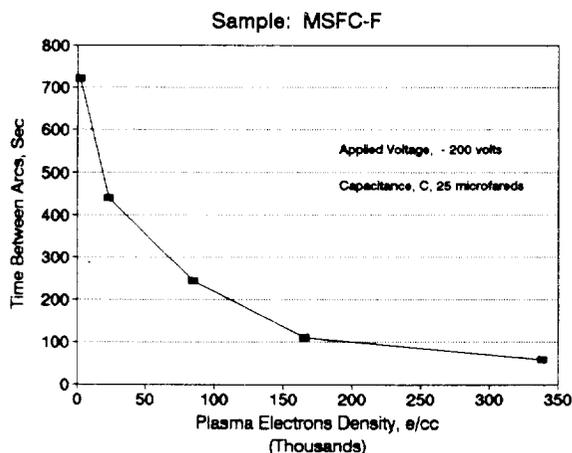
Figure 5. Average time between arcs as a function of plasma density for sample MSFC-F.

Corona Measurements

Corona inception voltages for various samples using metal electrodes are shown in Table IV together with the breakdown potential in the plasma experiments. It can be seen that the corona inception voltages followed the plasma breakdown potentials in that the higher the corona inception voltage the higher the breakdown voltage. Many of the samples, however, are grouped around a corona inception voltage of near 250 volts rms. (350 volts peak) although their breakdown values varied. What can be said is that samples that showed a corona

inception voltage of 200 volts and higher had plasma breakdown voltages of greater than 300 volts. This indicates a possible screening method for testing of anodization breakdown strengths.

Corona measurements using a liquid electrode, salt water, to simulate plasma contact gave unusual results. Corona initiation for all samples started at extremely low voltages, a few volts, and in

Table IV Corona Inception Voltage

Sample	Corona Inception Voltage, V(rms)	Breakdown Inception Voltage, V
MSFC-1	38	120
MSFC-F	66	200
LeRC-1	270	600
LeRC-2	250	400
LeRC-3	230	300
SSF-M1	270	600

some cases no measurements were possible because a short circuit appeared as the liquid made direct contact. A possible explanation of this may be as follows. In the process of manufacturing the anodized layer, a very thin solid "barrier layer" grows first, and then oxide columnar with open pores in between columns grows. (References 13-15.) The indicated thickness of oxide corresponds to the height of the columns and the barrier layer thickness. However, the true insulation thickness is the barrier layer only, which is much thinner. It is this layer that can easily break down. It is believed that the liquid electrolyte penetrates down to the thin barrier layer whereas the plasma sheath prevents the plasma from penetrating.

CONCLUSIONS

Anodized films vary greatly in their ability to withstand applied voltages, varying in breakdown strength from 100 volts to 2000 volts in the films tested here. Therefore it is possible to choose anodized films that will not breakdown under the potentials imposed on the space station structure, but the relation between the desired optical and thermal properties and the desired breakdown strength has not been established. The second type of arcing, that initiating at exposed metal sites, is not a factor since the inception voltage is above 200 volts and usually much higher.

Maximum arc current depends on the size of the capacitor used and the applied voltage, that is on the total capacitor charge. The maximum current varied approximately to the one-third power with the capacitance or total charge.

Experiments at low plasma density show that arcing that are initiated at bare metal sites occur at higher voltages as number density is lowered. Arcs triggered by insulation breakdown are independent of density. This confirms that true vacuums arcs are occurring in these ground tests, and the ambient plasma only as a trigger mechanism. For space plasma conditions where the metallic anode is not present some of the unusual phenomena of vacuum arcs will not appear and a type of arc not seen in ground experiments may occur.

Corona tests on anodized samples show that corona inception follows breakdown voltage, and can be used to screen new samples.

REFERENCES

1. Stevens, N. J., Berkopec, F. D., Grier, N. T., and Staskus, J., "Investigation of High-Voltage Spacecraft System Interactions with Plasma Environments," AIAA Paper 78-672, 1978.

2. Katz, I., Mandell, M. J., Schnulle, G. W., Parks, D. E., and Steen, P. G., "Plasma Collection by High-Voltage Spacecraft at Low Earth Orbit," Journal of Spacecraft and Rockets, Vol. 18, Jan-Feb. 1981, p. 79.

3. Grier, N., "Experimental Results on Plasma Interactions with Large Surfaces at High Voltages," NASA TM-81423, 1980.

4. N. J. Stevens, "Interactions between Large Space Power Systems and Low-Earth Orbit Plasmas," in Spacecraft Environmental Technology 1983, (C. K. Purvis and C. P. Pike, eds). NASA CP-2359 and AFGL-TR-85-0018, March 1985.

5. Kennerud, K. L., "High Voltage Solar Array Experiments," NADA CR-121280, March 1974.

6. Stillwell, R. P., Robinson, R. S., Kaufman, H. R., "Current Collection from the Space Plasma through Defects in high Voltage Solar Array Insulation," J. Spacecraft and Rockets 22, 1985, pp. 631-641.

7. Thiemann, H. and Bogus, K., "High-Voltage Solar Cell Modules in Simulated Low-Earth-Orbit Plasma," J. Spacecraft and Rockets, Vol 25, No. 4, July-August, 1988.

8. Mandell, J. J. and Katz, I., "Potentials in a Plasma Over a Biased Pinhole," IEEE Transactions on Nuclear Science, Vol. NS-30, 1983, pp. 4307-4310.

9. Vaughn, J., "SSF Chromic Acid Anodized Aluminum Arc Tests," Electrical Grounding Tiger Team Meeting, Huntsville, A., May 14-17, 1991.

10. Vaughn, J. "Experimental Investigation of Current Flow Dring Dielectric Breakdown," Electrical Grounding Tiger Team Meeting, Huntsville, Al, May 14-17, 1991.

11. Synder, D. B., "Characteristics of Arc Currents on a Negatively Biased Solar Array," NASA TM-83728, June 1984.

12. Jongeward, G. A., Katz, I., Mandell, M. J., and Parks, D. E.,"The Role of Unneutralzed Surface Ions in Nagative Potential Arcing," presented at chr 1985 IEEE 22nd Annual Conference, NSRE, Monterey, CA, July 22-24, 1985.

13. Symposium on Anodized Aluminum, Cleveland, OHio 1965. American Society for Testing Materials 1965, ASTM Sp Pub. 388.

14. Young, L., "Anodic Oxide Films," Academic Press, New York, 1961.

15. Lacey, E. A. and Sonnino, C., "Testing of Anodic Coatings on Aluminum alloys: A Review," *Testing of Metalic and Inorganic Coatings*, ASTM STP 947, W. B. Harding and G. A. Di Bari, Eds., American Society for Testing and Materials, Philadelphia, 1987, pp. 150-166.

Theoretical models of Kapton heating in solar array geometries

Thomas L. Morton
Sverdrup Technology
Lewis Research Center Group
Cleveland, OH 44135

ABSTRACT

In an effort to understand pyrolysis of Kapton in solar arrays, a computational heat transfer program was developed. This model allows for the different materials and widely divergent length scales of the problem. The present status of the calculation indicates that thin copper traces surrounded by Kapton and carrying large currents can show large temperature increases, but the other configurations seen on solar arrays have adequate heat sinks to prevent substantial heating of the Kapton. Electron currents from the ambient plasma can also contribute to heating of thin traces. Since Kapton is stable at temperatures as high as 600 Celsius, this indicates that it should be suitable for solar array applications. There are indications that the adhesive used in solar arrays may be a strong contributor to the pyrolysis problem seen in solar array vacuum chamber tests.

INTRODUCTION

During plasma compatibility tests of the SSF solar array blanket, pyrolysis of Kapton® (registered trademark of E.I. du Pont de Nemours & Co., Inc.) was seen.[1] While performing a dark test, the experimenters biased the array to a positive voltage, and read the current drawn from the ambient plasma. At +450 volts, a large spike was seen in the current. When the solar panels were removed from the test tank, a 1/8 inch hole was seen to have enlarged to about 1/2 inch diameter, with black char marks on the Kapton surrounding the hole.

Kapton was developed by du Pont in the early 1950's, and commercialized in 1966. It has been used extensively in the aerospace industry for 25 years with few problems. Various laboratory tests indicate that it does not pyrolyze until the temperature reaches at least 500° C.[2,3]

The objective of this paper is to determine possible causes of the pyrolysis, and determine the conditions under which they could be repeated. In addition, questions were raised about possible concerns when Kapton was used in the space environment. Under what conditions would Kapton pyrolyze, and would those conditions happen often?

METHOD OF CALCULATION

Cover slide (2 mils)

Silicon (8 mils)

Copper (1.4 mils)

Kapton (4.4 mils)
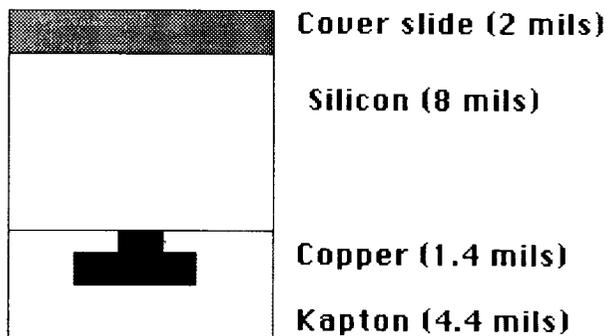
Figure 1 - Cross Section of Solar Cell

To study this problem, a computational heat transfer computer program was developed. A new program was needed because of the unique geometry of this problem. The solar cells are very thin, as shown in Figure 1. The entire thickness is about 15 mils, while the hole that was examined was 1/8 inch in diameter (125 mils). This disparity in scales does not allow the direction perpendicular to the surface to be treated in the same way as the directions parallel to the surface.

A schematic of the calculation geometry is shown in Figure 2. Both the hole, and the paths for heat gains and losses are shown. For most of the calculations, it was assumed that incident radiation was coming in

710

from a 300 K source. In addition, resistive heating was ignored for most of the calculations, since the system modeled was a dark solar panel, that is, no current was present.

Thus, the heat sources and paths are as follows:
1. Energy transfer from incoming electron, set equal to current times bias voltage.
2. Conduction of heat through copper, kapton, and silicon.
3. Radiation heat gain from external source at 300 K.
4. Radiation heat loss from surfaces, equal to $\epsilon\sigma T^4$, where ϵ is the emissivity , σ is the Stefan-Boltzmann constant, and T is the absolute temperature at the surface.
5. Internal Joule heating of copper from I^2R.
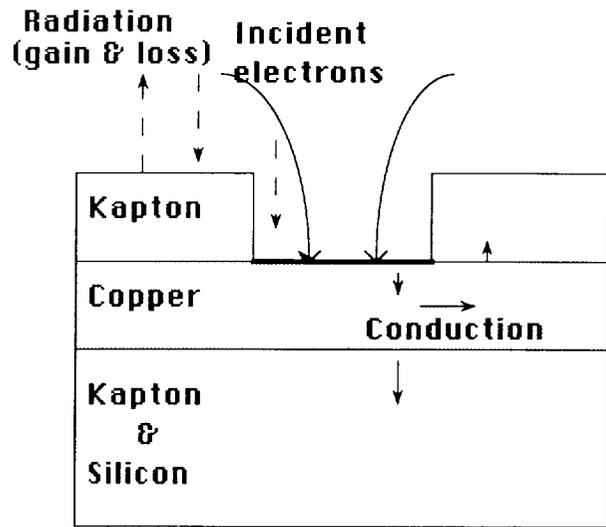
Different materials were included in this calculation, as listed in Table 1. The adhesives were treated as Kapton, since their thermal properties are similar, and could not be identified until late in the calculation.

Figure 2 - Heat pathways

Parameter \ Material	Kapton	Copper	Silicon
Density (g/cm³)	1.42	8.96	2.33
Specific Heat (J/g K)	1.09	.385	.702
Thermal Conductivity (W/cm K)	.00155	4.01	1.49
Resistivity ($\mu\Omega$ cm)	10^{18}	1.69x 10^{-8}	640.
Emissivity	.7	.65 (oxidized)	not used

The fundamental equation for heat transfer is

$$\rho c \frac{\partial T}{\partial t} = \kappa \nabla^2 T + \dot{Q} \qquad (1)$$

where ρ is the density, c is the heat capacity per unit mass, and κ is the thermal conductivity. Writing $\alpha = \kappa / (\rho c)$, the equations can be re-written as

$$\frac{\partial T}{\partial t} = \alpha \nabla^2 T + \frac{\dot{Q}}{\rho c} \qquad (2)$$

For this calculation, a finite difference approach was chosen. Boundaries between difference materials

were treated with a control volume approach. That is, the heat flowing into and out or a node from each direction would change the temperature of that node as if the node had the average properties of the materials around it.

In finite difference calculations, the time step is limited by stability requirements. If the calculation is done explicitly, it is necessary to maintain a ratio

$\frac{\kappa \Delta t}{\rho c (\Delta x)^2} < \frac{1}{2}$ where Δt is the time step, and Δx is the

distance scale. However, if the temperature at the succeeding time step depends implicitly on the surrounding temperatures, this restriction does not hold. Implicit calculations in one dimension are straightforward, but three dimensional calculations are difficult. For this reason, the calculation was done implicitly only in the z direction (i.e. perpendicular to

711

the surface of the Kapton).

For an explicit calculation, the temperature at time $n+1$ is simply

$$T_{i,j,k}^{n+1} - T_{i,j,k}^n + \frac{\alpha_x \Delta t}{(\Delta x)^2}(T_{i+1,j,k}^n - 2T_{i,j,k}^n + T_{i-1,j,k}^n)$$

$$\frac{\alpha_y \Delta t}{(\Delta y)^2}(T_{i,j+1,k}^n - 2T_{i,j,k}^n + T_{i,j-1,k}^n)$$

$$\frac{\alpha_z \Delta t}{(\Delta z)^2}(T_{i,j,k+1}^n - 2T_{i,j,k}^n + T_{i,j,k-1}^n) + \frac{\dot{Q}}{\rho c}$$

(3)

where $\alpha_i = \kappa_i/(\rho c)$, $i = x,y,z$. If we do an implicit calculation in the z direction, the formula becomes

$$T_{i,j,k}^{n+1} - T_{i,j,k}^n + \frac{\alpha_x \Delta t}{(\Delta x)^2}(T_{i+1,j,k}^n - 2T_{i,j,k}^n + T_{i-1,j,k}^n)$$

$$\frac{\alpha_y \Delta t}{(\Delta y)^2}(T_{i,j+1,k}^n - 2T_{i,j,k}^n + T_{i,j-1,k}^n)$$

$$\frac{\alpha_z \Delta t}{(\Delta z)^2}(T_{i,j,k+1}^{n+1} - 2T_{i,j,k}^{n+1} + T_{i,j,k-1}^{n+1}) + \frac{\dot{Q}}{\rho c}$$

(4)

Separating temperatures at time n from time $n+1$, we get

$$-d_z T_{i,j,k+1}^{n+1} + (1+d_z)T_{i,j,k}^{n+1} - d_z T_{i,j,k-1}^{n+1} = T_{i,j,k}^n + \frac{\dot{Q}}{\rho c}$$

$$+d_x(T_{i+1,j,k}^n - 2T_{i,j,k}^n + T_{i-1,j,k}^n)$$

$$+d_y(T_{i,j+1,k}^n - 2T_{i,j,k}^n + T_{i,j-1,k}^n)$$

(5)

where $d_i = \frac{\alpha_i \Delta t}{(\Delta i)^2}$ $i = x,y,z$. Von Neumann

stability analysis requires that $d_x < \frac{1}{2}$, $d_y < \frac{1}{2}$. There is no restriction on d_z.

Since d_i has the spatial increment in the denominator, it is very useful to not restrict d_z. For these calculations, a $\Delta t = .0001$ seconds was used. This keeps both d_x and d_y within the stability requirements.

The boundaries between different substances were

treated with a control-volume approach. That is, it was assumed that heat was transferred between nodes using an average of the transport properties of the materials involved. That is, the finite difference version of Eqn 1 becomes

$$(\rho c)_{av} \frac{(T_0^{n+1} - T_0^n)}{\Delta t} = \dot{Q} +$$

$$[\sum_{adjacent\ nodes} \kappa_{node} \frac{(T_{node}^n - T_0^n)}{(\Delta i)^2}]$$

(6)

where $(\rho c)_{av}$ is the average heat capacity per unit volume for the central node, κ_{ma} is the average thermal conductivity on the line joining an adjacent node to the central node, and Δi is the distance between the two nodes.

This method requires the boundaries between different substances to lie on nodes in the grid. However, this was rarely a strong restriction. The largest problem from that constraint was that the modeled hole through the Kapton was not circular. However, in light of the agreement with experimental results, the restriction was not found to be major.

TESTS OF THE CALCULATION

The model was tested in two different fashions. The original code was compared to analytic solutions to heat conduction problems, many of which were listed in *Carslaw & Jaeger (1959)*.[5] The results were compared to the calculation method to semi-infinite solids with a circular heat source, infinite slabs of different materials, and some cylindrically symmetric solutions.

By the time the code had been compared with the models, results from laboratory experiments were available, and the program could be compared to them.[6] The agreement with experiment was excellent (see Figure 3). The top of Figure 3 shows three temperature contours for a 1 inch copper strip, insulated by 1 mil of Kapton on each side. The hole is 1/4 inch in diameter, about 1/2 inch from the end of the copper trace. The middle contour is the experimental result, while the one on the right is the program. If the program includes the averaging of the infrared camera, and accounts for the lower emissivity of copper, the contour on the left is obtained. The graph on the bottom shows the temperature as a function of time for the one inch strip. Since computer time was limited, most runs were stopped at two

712

Figure 3
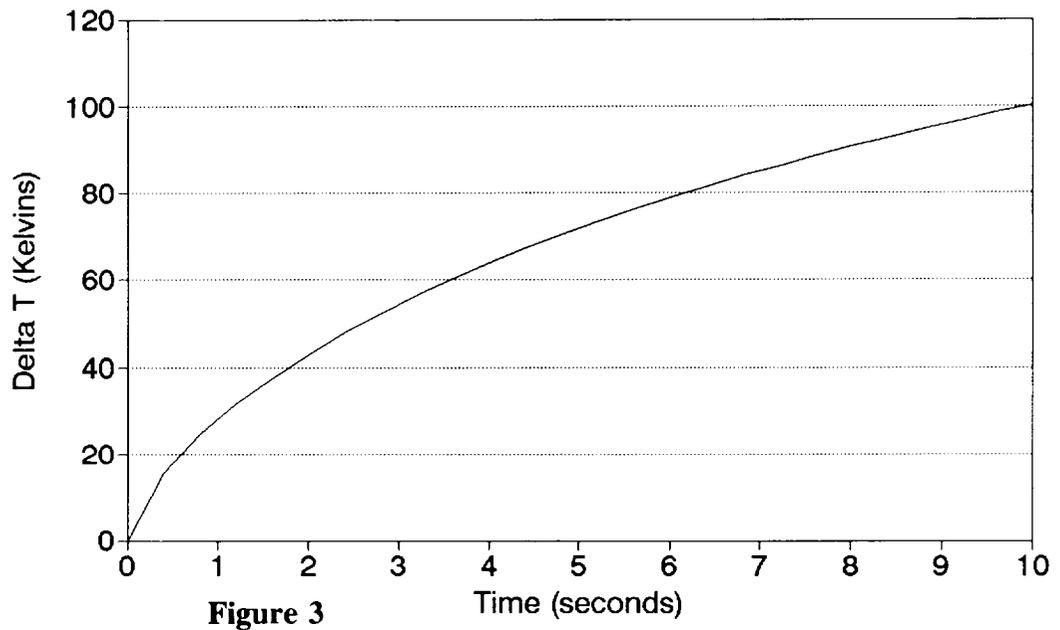
seconds of simulated time. The graph shows that the temperature at ten seconds is about two to three times the temperature at two seconds. It was therefore deemed appropriate to extend the calculation to models of solar cells.

The experiments covered a range of incident power from .005 W/cm² up to 13 W/cm². The computer program was able to duplicate the experimental temperature rise over the entire range of incident power. For the smallest copper strip tested, pyrolysis was seen at about 13 W/cm².

Two different examples of solar array geometries were examined. The first case modeled the connection between different silicon arrays, and consisted of a copper trace bounded on both top and bottom by Kapton. The second modeled the trace near the silicon, and included an electrical connection to the silicon as a possible pathway for heat to move. Since the original test was a dark

test, no resistive heating was included in these models. The results are shown in Table 2.

The first part of the Table shows the temperature rise for a solar cell copper trace, both theoretically after 2 seconds, and the experimentally after 10 seconds. The calculated temperatures should be multiplied by 2 to get an expected temperature rise for 10 seconds. The calculations show that even a small variation in hole size can change the expected temperature rise. In addition, three magnitudes of incident power are shown, to show how linear the problem is. The radiative heat term should generate a non-linearity, but that effect is quite small.

The second part of the Table shows the temperature rise expected for a copper trace near silicon. All three are after 10 seconds. The two calculations show the effect of a conductive pathway between the copper and the silicon. This is to model the weld joint which connects the copper to the silicon.

Table 2 - Model of thin trace embedded in Kapton

Incident power / area	1.4 mil x 5/32 inch 1/8 " hole	1.4 mil x 5/32 inch 5/64" hole	Experiment
Time	2 seconds		10 seconds
.1 W / cm²	.7867249 K	.4821346 K	.55 K
1 W / cm²	7.8614550 K	4.8192897 K	7.5 K
10 W / cm²	77.9610901 K	47.971920 K	80. K

Model of thin trace near Silicon

Incident power / area	Conductor from Copper to Silicon	No conductor from Copper to Silicon	Experiment
Time	10 seconds		
1 W / cm²	.8 K	1.1 K	2.2 K

NASCAP/LEO, a space-charge analyzing computer program, can predict the currents expected for a biased hole exposed to the ambient Low Earth Orbit (LEO) plasma. For 1/4 and 1/2 inch diameter holes, these I-V curves are shown in Figure 4 for possible voltages and densities expected on the solar array[7]. The highest plasma densities expected for LEO orbits would be during Solar Maximum conditions, when the density can reach 4 x 10¹² ions per cubic meter. At this density, the area of the solar array at 160 volts would collect 70 micro-amps of electron

current. The model predicts a temperature rise of less than 1 degree Celsius for this incident power. This is not enough to cause significant heating of the copper or Kapton. For comparison, the laboratory test that saw pyrolysis had about 1 amp at 450 volts.

During daylight, the solar arrays are generating current through the traces. If resistive heating in the copper trace is included, the Kapton achieves an equilibrium temperature rise of about 10 degrees Celsius. In addition, the sunlight heats the arrays to about 70

714

temperature rise of about 10 degrees Celsius. In addition, the sunlight heats the arrays to about 70 degrees Celsius. Even an extra 1 degree Celsius from plasma heating will still not achieve pyrolysis.

CONCLUSIONS

Pyrolysis of Kapton has been duplicated in a laboratory plasma tank for a solar array type of geometry.

The calculations confirm the experimental results that the change in temperature observed in these samples is proportional to the product of incident current times the bias voltage. In LEO conditions, it does not appear likely that Kapton will pyrolyze due to incident energy from the plasma at Space Station voltages.

ACKNOLOWEDGMENTS

REFERENCES:

1. Felder, Marian, private communications.

2. Jellinek, H.H.G., "Degradation and Stablization of Polymers", Elsevier, New York, 1983.

3. Punderson, J. O. and Heacock, J. F., "Why Long-term Performance Exceeds Some Limited Laboratory Projections", 34th International Wire and Cable Symposium, Cherry Hill, NJ, November 19-21, 1985

4. Jaluria, Yogesh and Kenneth E. Torrance, COMPUTATIONAL HEAT TRANSFER, Hemisphere Publishing Corporation, New York, 1986.

5. Carslaw, H.S. and J.C. Jaeger, CONDUCTION OF HEAT IN SOLIDS, Clarendon Press, Oxford, 1959.
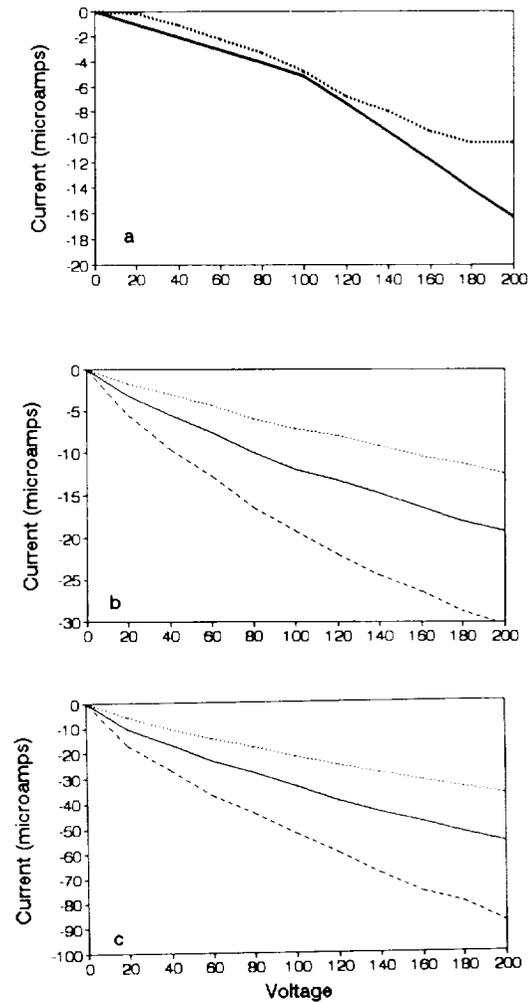
Figure 4 - (a) Dotted line is NASCAP/LEO, solid is lab result. (b) Current in space for densities of 10^6, 2×10^6, and 4×10^6 electrons/cc, 1/8 inch hole. (c) Current in space for densities of 10^6, 2×10^6, and 4×10^6 electrons/cc, 1/4 inch hole.

6. Grier, Norman T., "Experimental verification of Kapton Pyrolysis", SOAR-91, Houston, TX July, 1991.

7. Morton, T. and R. Chock, "NASCAP/LEO simulations of current collection by pinholes", (unpublished).

Sputtering of ions from Cu and Al by low energy oxygen ion bombardment

T. G. Eck, L.-Y. Chen, and R. V. Hoffman
Department of Physics, Case Western Reserve
University, Cleveland, OH 44106

We are investigating the sputtering of Cu^+ and Al^+ ions from copper and aluminum when the target surface is bombarded with low energy oxygen ions (20 to 500 eV). We hope to be able to determine both the ion yield as a function of oxygen ion energy and the kinetic energy distribution of the sputtered ions at a given oxygen ion energy.

Session E4: INTERACTION LABORATORY EXPERIMENTS -II

Session Chair: G. B. Hillard

717

LABORATORY STUDY OF THE TEMPORAL EVOLUTION OF THE CURRENT-VOLTAGE CHARACTERISTIC OF A PROBE IN THE WAKE OF AN OBJECT IMMERSED IN A PULSED FLOWING PLASMA

S. Meassick and C. Chan
Department of Electrical and Computer Engineering
Northeastern University, Boston, MA 02115

ABSTRACT

Measurements of the current-voltage characteristics of a Langmuir probe in the near wake of a disk immersed in a pulsed flowing plasma have been made. A 1 cm diameter biasable sphere was placed in the ion-free near wake region of a 10 cm diameter disk immersed in a Mach 8 pulsed flowing plasma. The current-voltage characteristic of the sphere was observed as a function of time as the sphere bias was scanned from $-5000\ V$ to $+1000\ V$. The collected current is found to be monotonically increasing with increasing positive bias voltage but exhibits a threshold voltage for current collection as the bias voltage becomes more negative. Potential measurements in the wake region were made for a sphere bias voltages below, at, and above the current collection threshold for a number of times during the wake formation period. The time evolution of the potential profile is shown to change as the sheath around the biased sphere is established. Predictions from the particle trajectory code SIMION are compared with data, showing excellent agreement in the prediction of the current collection threshold.

INTRODUCTION

The issue of current collection by a charged object in the wake of another, usually larger, object is becoming more important as larger platforms, especially in near polar orbits are contemplated. With structures that are many thousands of Debye lengths in size the plasma density in the near wake region will be reduced to a small fraction of the ambient plasma density. Objects in this wake region will not draw any appreciable current until the potential on the object is sufficient to draw charged particles across the depleted plasma region from the plasma flowing around the object creating the wake. In addition, in the case of ions, there may also be an angular momentum barrier to overcome for ion current collection.

Due to their importance, the study of plasma wakes has long been an area of interest in space physics. Numerous investigators, including *Samir et.al.* [1979, 1981, and 1986] and *Medved* [1969], have investigated problems of the environment of a small satellite in the ionosphere. The wake region of large objects such as the space shuttle has been investigated by *Riatt et.al.* [1987] and *Murphy et.al.* [1986]. In fact in 1985 the *Plasma Diagnostics Package* was flown as part of the Spacelab-2 mission. This free flying satellite was designed to make comprehensive measurements of the plasma density, temperature and turbulence in the near, mid and far wake of the space shuttle.

The problem of current collection in the wake of an object in space has received considerably less attention. But for some theoretical predictions of a threshold voltage for ion current collection and some laboratory simulations (*Chan et.al.* [1989]) that have verified these predictions, very little work has been done in this area. In order to gain a fuller understanding of the problem of a Langmuir probe in the wake of a conducting body (ie. current collection in the wake of another body) we have done a series of laboratory experiments to study the dynamics of the current-voltage characteristics of a biased Langmuir probe placed in the wake of a conducting body. We have made comprehensive measurements of the current collected by and the plasma potential around a Langmuir probe as the wake region is established.

DESCRIPTION OF EXPERIMENT

The pulsed plasma wind experiment (PPWE) [*Measssick et.al.*, 1991; *Chan et.al.*, 1984; *Morgan et.al.*, 1987] that we have developed to study the dynamics of current collection of a Langmuir probe in the wake of another body in low earth orbit (LEO) is shown schematically in figure 1. This device consists of a source region and a flowing plasma region. The source region is 30 cm in length and 50 cm in diameter while the flowing plasma region is 70 cm in length and also 50 cm in diameter. The plasma is created in the source region via a multidipole discharge and is allowed to flow into the flowing plasma region where the object creating the wake structure and the Langmuir probe are

present. The energy of the ions is adjustable through a bias voltage applied to the source region. The dynamics of the current collection problem are studied by pulsing the plasma into the flowing plasma region and measuring the response of the current-voltage characteristic of a Langmuir probe and the plasma potential profiles around it.

This experimental arrangement has allowed us to create a pulsed flowing plasma with a variable ion Mach number in the range of 1 to 10. For the experiments reported on in this paper, the plasma flowing into the flowing plasma region has a maximum density of 1×10^8 cm^3 and a parallel and perpendicular ion temperature of 4 eV and 2 eV respectively. The electrons have a perpendicular temperature of 2 eV and a parallel energy of approximately 6 eV. The plasma expanding into the flowing plasma region is not initially in contact with either the side or end walls of the chamber and does not reach the end or side walls of the chamber for approximately 120 μs after the plasma is pulsed into the flowing plasma region.

In order to measure the dynamics of current collection of a Langmuir probe in the wake of a larger object a grounded 10 cm diameter conducting disk was placed 10 cm downstream of the source region. This disk was grounded in order to have constant boundary conditions and eliminate the large floating potential fluctuations of a disk placed in a pulsed flowing plasma as seen by *Morgan et.al.* [1989]. A spherical Langmuir probe, 1 cm diameter, was placed on axis 5 cm downstream of the disk. The Langmuir probe was biasable through an external power supply from +5000 to -5000 V. For these experiments, the base pressure of the device was approximately 5×10^{-7} Torr while the operating pressure during measurements was 5×10^{-5} Torr of Argon.

Collecting and emitting Langmuir probes were utilized to monitor the plasma parameters in the source and flowing plasma regions. Collecting Langmuir probes, consisting of 0.64 cm diameter disks were used to monitor the plasma density and the electron temperature in the source region. Retarding Potential Analyzers (RPA) were used to measure the ion temperature in the source region and to measure the ion and electron distributions in the flowing plasma regions. Emissive Langmuir probes were used to measure the plasma potential. These probes were operated in the limit of zero emission utilizing the inflection-point method (*Smith et.al.* [1979]) in order to minimize the perturbation of the plasma. All probes were constructed to have a maximum time resolution of 1 μs and were scannable to cover the source and flowing plasma regions.

Data from all of the probes were sampled with a boxcar averager that was triggered when the plasma began flowing into the flowing plasma region. This allosed sampling of the data at the times of interest and reduction of noise.

EXPERIMENTAL RESULTS

Measurements of the current collected by the Langmuir probe were made over a bias voltage range of +500 to -5000 V. Figure 2a shows the amount of current collected by the probe as a function probe bias voltage at 100 μs after the start of the plasma pulse. At this time the wake region is already well established and has reached steady-state conditions. It is evident that for positive bias voltages, there is a rapid increase in the current, reaching a saturation current of somewhat greater than 1 mA at a bias voltage of approximately 100 V. For negative bias voltages applied to the Langmuir probe no current is collected until the bias voltage reached approximately -2.2 kV after which ion current is collected.

Figure 2b shows the current collected by the Langmuir probe as a function of time. It is evident that the threshold voltage for ion current collection increases rapidly at early times. At late times (100 μs) the threshold for ion collection is approximately -2.2 kV and increases to approximately -5 kV at very early times (30 μs) as the wake is just being established.

In order to gain an understanding of the temporal evolution of the voltage-current characteristic of the Langmuir probe, simultaneous measurements of the plasma potential throughout the near and mid wake region were made. These potentials measurements were made with Langmuir probe bias potentials of -2000, and -3000 V. These bias voltages were chosen to gain an understanding of the current collection threshold characteristics. The dynamics of the potential profiles were studied by taking measurements of the plasma potential during the formation of the wake from the time that the flowing plasma just arrives in the wake region (30 μs) until a steady state has been achieved (100 μs).

Figure 3a-d shows the temporal evolution of the potential profiles just below the current collection threshold at -2000 V. From part a of this figure it is evident that the negative potential profile extends far past the disk radius at 30 μs (the -10 V equal potential profile extends out to a radius of 8 cm). There are negative potential "wings" that extend upstream of the disk. At 40 μs, shown in part b of this figure the -10 V equal potential profile has decreased in radius so that it now only extents slightly past the disk radius. In addition the axial extent of the -10 V equal potential profile downstream of the Langmuir probe has decreased from 12 to 8 cm and the negative potential "wings" have disappeared. In part c of the figure, at 80 μs, the -10 V equal potential profile extends to a radius of less than that of the disk, indicating that the large electric field region is now confined to small radii. It is also evident that at this time that an enhanced potential

region is forming on axis downstream of the sphere (for axial distances greater than 10 cm). Figure 3d shows that for late times all of the negative potential regions are at radii less than the disk radius. In addition, the enhanced potential region first seen in at 80 μs is now much more pronounced.

The temporal evolution of the plasma potential for a Langmuir probe bias voltage above the current collection threshold voltage is shown in figure 4a-d. With a Langmuir probe bias potential of -3000 V the sheath region extends far from the probe. Even at 100 μs the -10 V equal potential contour is at a radius of 5 cm, the same as the radius of the disk creating the wake. In addition, the on axis potential enhancement seen with a Langmuir probe bias of -2000 V is no longer evident. The "wings" that only extended upstream of the disk at 30 μs for lower probe bias voltages now are also present at 40 μs.

DISCUSSION

The two most significant features of the current-voltage characteristic of a biased Langmuir probe placed in the wake are that there is a threshold voltage for ion current collection and that this threshold voltage is a strong function of time during the formative stages of the wake.

The threshold voltage for current collection has been seen in steady state experiments performed by *Chan et.al.* [1989]. In those experiments a 10 cm diameter disk was placed in a 1.7 by 1.7 m vacuum chamber in order to eliminate wall effects. The threshold voltage for ion current collection with a 1 cm diameter Langmuir probe 5 cm downstream of the disk is -2200 V, almost identical to the threshold voltage reported in this experiments for late times. The measured potential profiles are nearly identical to the ones presented in this paper for late times.

In order to gain some insight into the temporal variations of the current collection threshold we have utilized the charged particle trajectory code SIMION (*Dahl and Delmor* [1988]). This is a two dimensional Poisson solver that solves for space potentials, not including space charge effects, given boundary conditions and allows the tracking of charged particle trajectories. In order to utilize this code the measured potential profiles were entered and particle trajectories were followed.

Figure 5a,b shows the trajectories for ions with Langmuir probe bias potentials of -2000, and -3000 V respectively. The angle of the ion trajectories is due to the finite perpendicular energy of the ions. From these figures it is evident that ions are only collected by the sphere for bias potentials greater than -3000 V. For bias potentials smaller than -3000 V the ions are deflected by the potential gradients with their trajectories crossing 10 to 15 cm downstream of

the disk. This is at the same axial location where the is an enhancement of the plasma potential.

Figure 6a-c shows the ion trajectories for the Langmuir probe biased at -3000 V at 40, 80, and 100 μs respectively. From these figures it is evident that for early times, where the sheath region extends far from the Langmuir probe, the ion orbits do not hit the probe. As the sheath region shrinks the ions come closer to the probe until their orbits intersect the probe and are collected. In addition, at earlier times there is very little focusing of the ion trajectories downstream of the probe (there is no enhanced potential region downstream for early times).

From this simulations it is evident that the reason that the current collection threshold is much larger early in time is that the sheath region is large. When the sheath region is large the electric field is not large enough so that the particle orbits have a minimum approach distance to the Langmuir probe so that they will be collected.

CONCLUSIONS

The temporal evolution of the current collection of a Langmuir probe inserted in the near-wake region of a conducting body has been studies in the laboratory with the aid of a pulsed plasma wind experiment. This experimental arrangement allowed the study of the dynamics of current collection without the influence of the vacuum chamber walls. Measurements of the temporal evolution of the current-voltage characteristic of a Langmuir probe in addition to direct probe measurements of the plasma potential during the formative process of the wake region were made. From these experiments it was determined that a threshold voltage exists for the collection of ions by a biased Langmuir probe in the wake region. This threshold voltage decreases as the sheath region around the probe is established and the electric field increases. No similar threshold voltage for the collection of electron current exists due to the small electron mass.

REFERENCES

1. Chan, C., Browning, J.,Meassick, S., Morgan, M., Cooke, D., Enloe, D., and Tautz, M., "Current Collection in a Spacecraft Wake; Laboratory and Computer Simulations", Proceedings of the Spacecraft Charging Technology Conference, Monterey, CA, 31 October-3 November, 1989.

2. Chan, C., Hershkowitz, N., Ferreira, A., Intrator, T., Nelson, B., and Lonngren, K., "Experimental ob-

servation of self similar plasma expansion", PHYS. FLUIDS, Vol. 2, 1984, pp. 266.

3. Dahl, D. A., and Delmor, J. E., "SIMION", EGG CS-7233 Rev. 2, EG+G Idaho Inc., Idaho Falls, April 1988.

4. Meassick, S., Chan, C. , Qian, Y., Sroda, T., Azar, T., and Allen, R., "Temporal Study of Wake Formation Behind a Conducting Body", J. GEOPHYS. RES., (Accepted for publication).

5. Medved, D. B., "Measurement of ion wakes and body effects with the Gemini/Agena satellite", RAREFIED GAS DYN., Vol. 1, 1969, pp. 1525.

6. Morgan, M. A., Chan, C. , Cooke, D. , and Tautz, M., "The dynamics of charged particles in the near wake of a very negatively charged body laboratory experiment and numerical simulation", IEEE TRANS. PLASMA SCI., Vol. 17, 1989, pp. 220.

7. Morgan, M. A., Chan, C., and Allen, R., "A laboratory study of the electron temperature in the near wake of a conducting body", GEOPHYS. RES. LETT., Vol. 14, 1987, pp. 1170.

8. Murphy, G. B., D'Angelo, N., Kurth, W. S., and Pickett, J., "Measurement of plasma parameters in the vicinity of the space shuttle", PLANET. SPACE SCI., Vol. 34, 1986, pp. 993.

9. Raitt, W. J., Eccles, J. V., Thompson, D. C., Banks, P. M., Williamson, P. R., and Bush, R. I., "Plasma parameters in the near wake of the space shuttle", GEOPHYS. RES. LETT., Vol. 14, 1987, pp. 359.

10. Samir, U., Gordon, R. , Brace, L., and Theis, R., "The near wake structure of the atmospheric explorer C (AE C) satellite: A parametric investigation", J. GEOPHYS. RES., Vol. 84, 1979, pp. 513.

11. Samir, U., and Fontheim, E. G., "Comparison of theory and in situ observations for electron and ion distributions near the wake of the Explorer 31 and AE C satellites", PLANET. SPACE SCI., Vol. 29, 1981, pp. 975.

12. Samir, U., Wildman, P. J., Rich, F., Brinton, H. C., and Sagalyn, R. C., "About the parametric interplay between ionic mach number, body size, and satellite potential in determining the ion depletion in the wake of the S3 2 satellite", J. GEOPHYS. RES., Vol. 86, 1981, pp. 11 and 161.

13. Samir, U., Comfort, R. H., Chappell, C. R., and Stone, N., "Observations of low energy ions in the wake of a magnetospheric satellite", J. GEOPHYS. RES., Vol.91, 1986, pp. 5725.

14. Smith, J., Hershkowitz, N., and Coakley, P., "Inflection point method of interpreting emissive probe characteristics", REV. SCI. INSTRUM., Vol. 50, 1979, pp. 210.
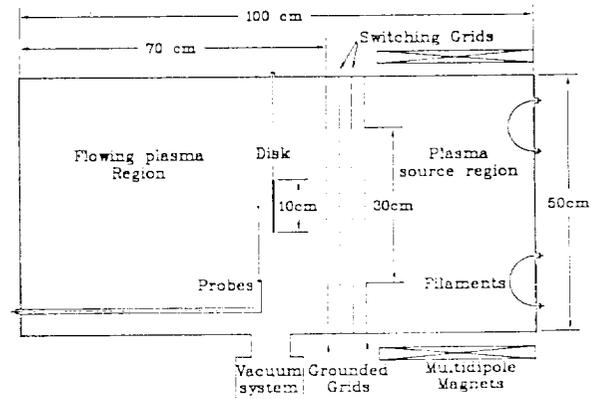
Figure 1: Schematic diagram of the Pulsed Plasma Wind Experiment (PPWE).
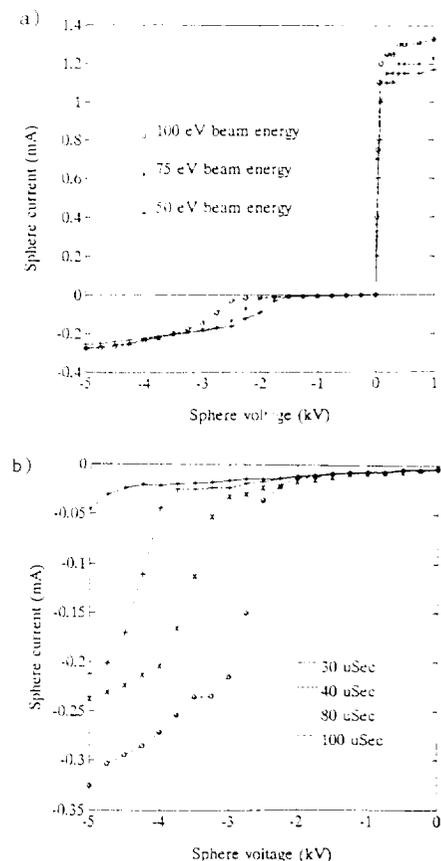
Figure 2: (a) The current-voltage characteristic of a Langmuir probe placed in the near wake region of a conducting sphere. The wake is created by a Mach=8 plasma flow. (b) The temporal evolution of the current-voltage characteristic of the Langmuir probe during the formation of the wake.

721
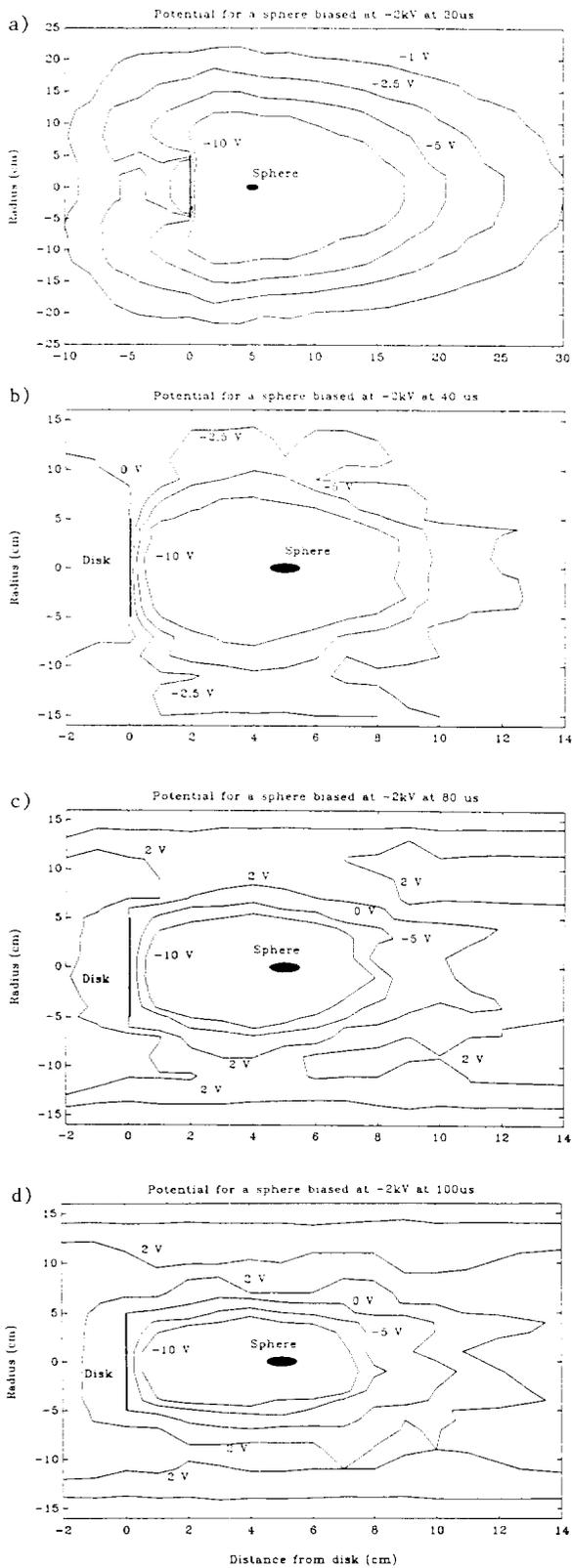
Figure 3: Plasma potential profiles around a Langmuir probe biased at -2000 V at (a) 30 μs, (b) 40 μs, (c) 80 μs, and (d) 100 μs.
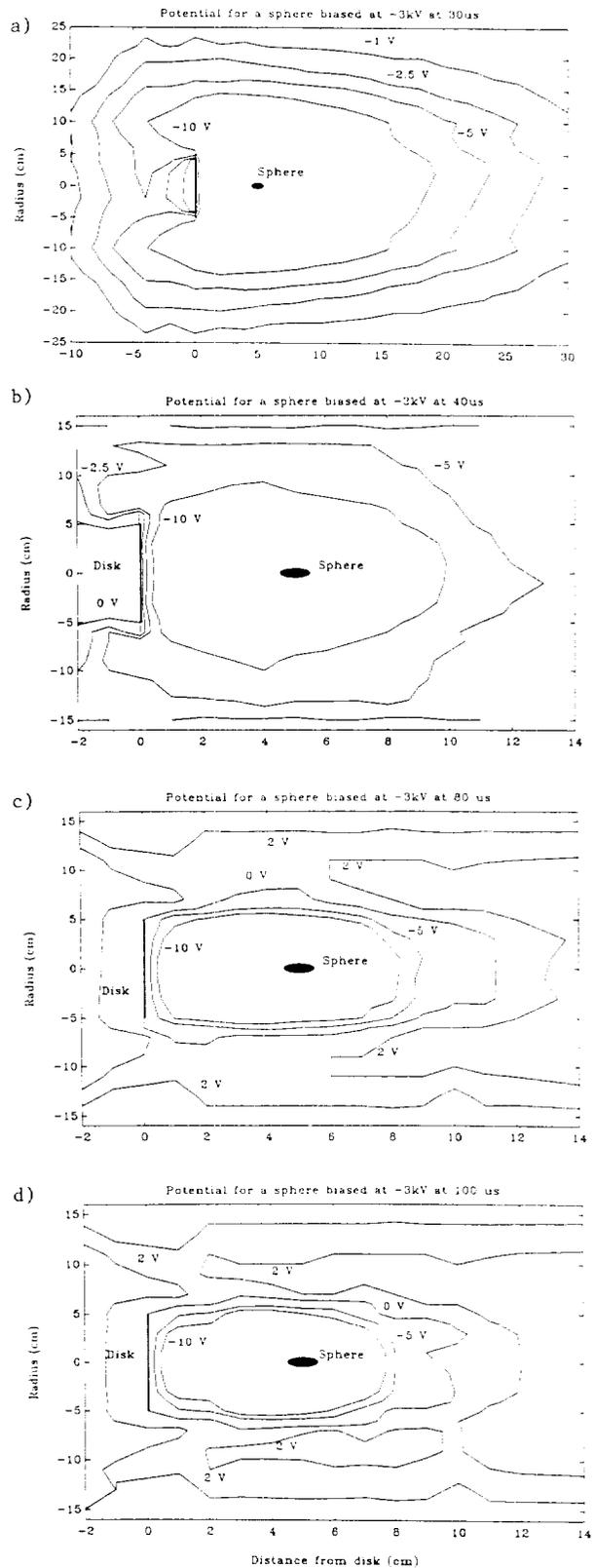
Figure 4: Plasma potential profiles around a Langmuir probe biased at -3000 V at (a) 30 μs, (b) 40 μs, (c) 80 μs, and (d) 100 μs.
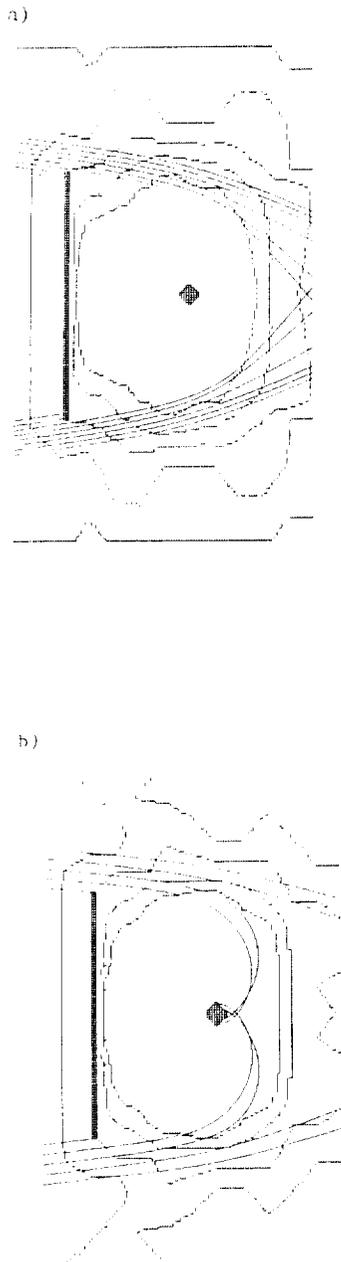
Figure 5: Ion trajectories predicted with SIMION for Langmuir probe bias potentials of (a) -2000 V and (b) -3000 V show that ions are only collected for -3000 V.
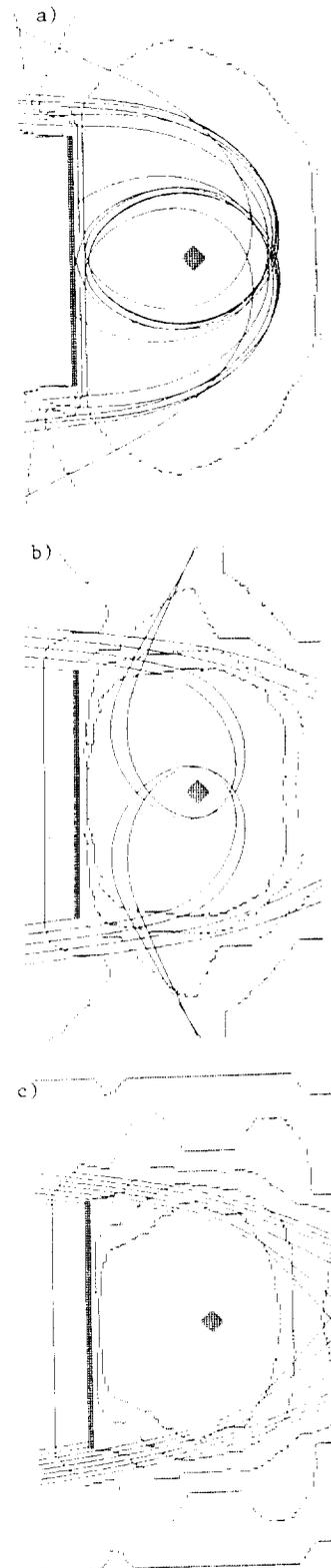
Figure 6: Ion trajectories predicted with SIMION for a Langmuir probe bias potential at -3000 V at (a) 40 μs, (b) 80 μs, and (c) 100 μs.

CURRENT FLOW IN A PLASMA CAUSED BY
DIELECTRIC BREAKDOWN

J. A. Vaughn and M. R. Carruth, Jr.
George C. Marshall Space Flight Center
MSFC, AL 35812

and

P. A. Gray
Sverdrup Technology
Huntsville, AL 35812

ABSTRACT

Spacecraft with a thin dielectric coating on the outer surface of the structure which are biased ($^-$ 200 V) negative relative to the atmospheric plasma are susceptible to dielectric breakdown. This paper will present experimental tests designed to measure the electron current flow from the structure through the plasma during the arc. The current path was examined in three parts: the electrons supplied through the structure and the arc to the outer structure, the expansion of the arc into the ambient plasma and the return current through the ambient plasma. The measured electron current either flowing from the plasma or supplied to the plasma by the arc in each case was compared to the random thermal electron current which could be collected. The results of the tests show a spacecraft is capable of supporting arcs with peak currents greater than thermal electron currents, and these currents will be dependent upon the amount of stored charge in the structure (i.e. the structure's surface area and dielectric thickness). Also, the results of these tests show that it is possible for structures with a self capacitance of 10 microFarads to see peak currents of 90 A and structures with 1000 microFarads (i.e. capacitance of one Space Station Freedom module) to produce peak currents of 1000 A.

INTRODUCTION

All spacecraft to date, including communications satellites, interstellar probes, and the proposed space station, have or will use solar arrays to harness the sun's energy for power. Most of the solar arrays used in the past have been of the low power variety, generating voltages near 30 volts. One exception to this was Skylab which flew a 90 V solar array [1].

As spacecraft become larger and more complex they need more power to operate. In an effort to increase the power available to the spacecraft, the voltage of the solar arrays had to be increased to prevent the need for developing electrical distribution systems capable of handling large currents. This has placed these new larger spacecraft in a regime where the physical interactions between the high voltage solar array [2-5] and the plasma have become significant.

A generating solar array in low earth orbit (LEO) must come to an equilibrium point such that it collects an equal number of ions and electrons. Because electrons are more mobile than ions and easier to collect, the potential of the solar array will be split to collect equal numbers of ions and electrons. This means that a large portion of the solar array will be negative of the ambient plasma, which is the ultimate reference point, in order collect the needed number of ions. Laboratory and flight data have shown this to be the case [6-9]. Because the common practice is to ground or reference the spacecraft to the negative side of the solar array, the potential of the spacecraft will be at the same potential relative to the plasma as the negative side of the solar array. For these new higher powered spacecraft, this places the spacecraft at a high voltage negative of the ambient plasma potential. In the case of Space Station Freedom the structure will

724

be approximately 140 V negative of plasma potential [10].

Some spacecraft have been designed with a thin dielectric coating on the outer surface to protect the spacecraft from the LEO environment. For Space Station Freedom the selected coating is anodized aluminum with an anodic thickness of 0.05 mils to 0.2 mils. A spacecraft immersed in the plasma with this thin dielectric material will build up charge like a large capacitor. The electrons which are collected by the solar array are stored near the surface of the anodic film and the ions from the ambient plasma will be collected on the space side of the anodic film. If the dielectric strength of the coating is not sufficient to withstand the applied voltage stress, dielectric breakdown will occur. Dielectric breakdown has been observed in the laboratory at applied voltages of -80 V for anodized coatings with a thickness of 0.05 mils and -120 V for 0.1 mil thick anodized aluminum [10].

Dielectric breakdown in a conductive plasma has been shown to produce large currents in a short time [10]. The peak currents observed in the laboratory were 90 A at 10 microFarads and 1000 A at 0.1 Farads. In order for currents of this magnitude to be possible the ambient plasma and the spacecraft in question must interact to maintain charge neutrality. If the spacecraft dumps large amounts of electrons into the plasma, the ions that built up on the surface must be neutralized to maintain charge neutrality. This paper will present experimental results describing the current paths needed to neutralize the ions on the surface to support these large currents.

The tests were separated into three parts, which are shown in Fig. 1, to examine the entire current flow path. The current path begins by forming a dense arc plasma by vaporizing and ionizing the aluminum metal that expands into the ambient plasma. The expansion rate of the arc plasma has been measured and it's significance will be discussed. Also, evidence of the arc plasma's ability to discharge the anodized surface will be presented.

Remote areas of the spacecraft structure, which are isolated from the arc plasma will not be able to interact with the expanding plasma. The ambient plasma will only see the structure undergo a significant positive voltage transient during the arc. The ability of this voltage transient to extract electrons from the ambient plasma has been examined, and experimental data will show the current to the surface of the structure to be much
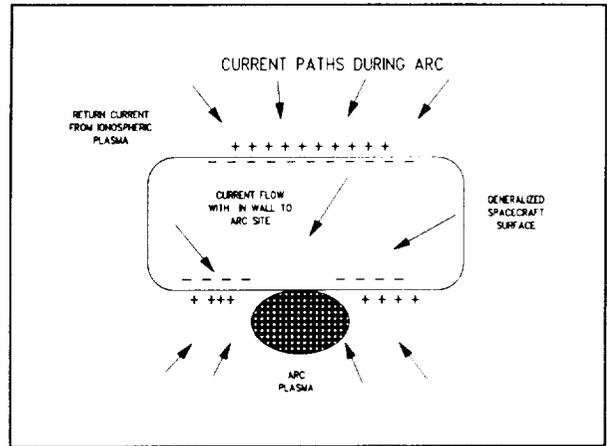
Figure 1 Example of Electron Flow in the Plasma

higher than thermal electron current collection.

EXPERIMENTAL TEST APPARATUS

The tests were performed in a 4 ft diameter by 6 ft long cylindrical vacuum chamber capable of a base pressure in the high 10^{-8} torr range. An argon plasma was produced in the chamber using a hollow cathode plasma source [11]. A hollow cathode plasma source was chosen for these tests because of its ability to produce a fairly uniform unperturbed plasma. Because hollow cathode plasma sources require a flow of gas to operate, the pressure during operation of the plasma source was in the low 10^{-4} torr range.

The spacecraft surface was simulated in these tests with chromic acid anodized aluminum plates with an anodic thickness layer ranging from 0.05 mils to 0.10 mils. This material was chosen because of the direct implication it has as the baseline material for Space Station Freedom. The back side of the 7.6 cm by 12.7 cm samples were coated with kapton to insulate them from the plasma. By insulating the back side of the samples from the plasma, it was easier to understand how the ambient plasma was interacting with the samples. The samples were biased negative with a DC power supply which was current isolated from the samples by a 10 K ohm resistor.

Schematics of the various test configurations are shown in Figs. 2,3 and 4. All test configurations were identical in chamber size, type of plasma source, and bias supply. Below, is a description of the variations in each test configuration.

Arc Expansion into the Ambient Plasma

Figure 2 shows the test configuration used to measure the arc plasma expansion rate. A single anodized aluminum plate was

725

placed in the center of the chamber and biased 160 V negative of ground and allowed
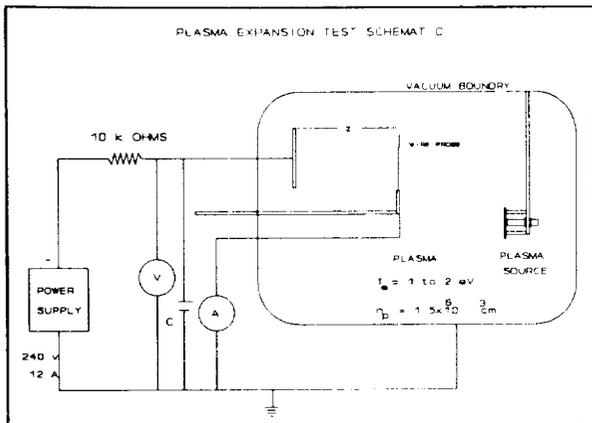
Figure 2   Arc Plasma Expansion Test Schematic

to arc. Because the plates themselves did not have sufficient charge capability, an 8 microFarad capacitor was placed in the circuit to provide a small amount of charge. A 0.025 cm (0.01 in) diameter and 12.7 cm (5 in) long cylindrical Langmuir probe which could be moved axially was placed into the chamber. The probe was grounded through an ammeter to measure the current surge from the arc. One channel of an oscilloscope was connected to the ammeter to measure the current surge as a function of time and the other to the voltage bias on the plate. The oscilloscope was triggered off the voltage bias on the plate to signal the start of the arc.

Current and voltage data were recorded at 5 cm increments over a 50 cm range. The data collected were analyzed in terms of the time for the current surge to reach a given probe position.

Current Supplied by the Structure

Tests to examine the current flow through the structure during an arc were performed (Fig. 3) by placing two anodized aluminum samples of identical size and anodization condition approximately 36 cm (14 in) apart in the vacuum chamber. Each plate was isolated from the other as well as the tank wall. The two plates were connected to each other outside the tank and an inductive current probe was placed around the line between them to measure current transferred during an arc. The arrow shown in the figure indicates the direction of positive electron current. Both plates were biased to the same potential in the tank and allowed to arc. One channel of an oscilloscope was connected to the current probe to measure current transients between the two plates
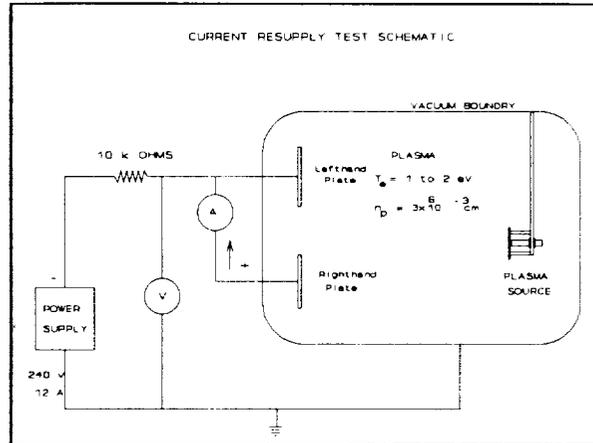
Figure 3   Schematic of Two Plate Experiments

and the other to monitor the voltage bias. The oscilloscope was triggered from the plate voltage to signal the start of the arc. The current and voltage data were then numerically analyzed to measure the total integrated charge transferred to the other plate during each arc.

Return Current from the Plasma

A single anodized aluminum plate was placed in the center of the chamber and biased -140 V to -160 V. In this set
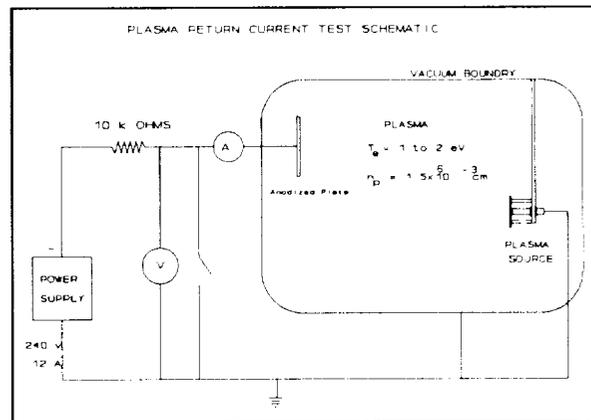
Figure 4   Return Current from the Plasma Test Schematic

oftests the plate shown in Fig. 4 was not allowed to arc but was instantaneously grounded using a switch to simulate the voltage transient the structure would see if an arc occurred. An ammeter was connected to the plate, which has a known capacitance, to measure the amount of electron current caused by the neutralization of ions by the ambient plasma electrons. In this set of tests the plasma source which was allowed to float in all other tests, was grounded to simulate the infinite well of electrons one would see in space. One channel of an

726

oscilloscope was connected to the ammeter the other to the plate bias. The oscilloscope was triggered off the plate bias. Once the voltage transient was applied, the current data were recorded as a function of time. The recorded current data were numerically integrated and the amount of charge flowing calculated.

EXPERIMENTAL RESULTS

After the initiation of an arc by dielectric breakdown, electron current flows from the structure to the ambient plasma while the ambient plasma is returning electrons to the structure. The flow of electrons through the plasma begins with the production of a dense arc plasma
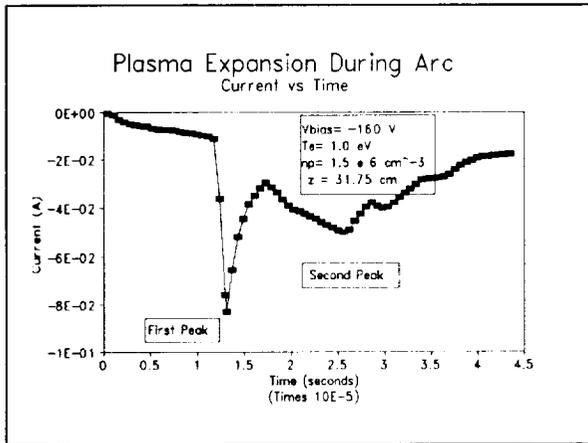
Figure 5   Typical Arc Plasma Expansion Data

by vaporization and ionization of the bulk aluminum metal. The electrons stored in the structure which is negative of the ambient plasma provide the necessary energy to produce a plasma during the arc. The arc plasma then expands into the ambient plasma in all directions neutralizing the part of the structure it contacts.

Typical data collected while measuring the expansion rate of the arc plasma are shown in Fig. 5. The current trace presented in this figure was recorded with the Langmuir probe approximately 32 cm from the anodized aluminum sample. All current traces recorded showed the two peak structure observed in Fig. 5. This figure shows an intense plasma wave front passes the probe first followed by a second less intense wave. The intensity of both fronts (i.e. 80 mA and 50 mA) is greater than the thermal electron current (0.1 micro Amps) that could be collected by the probe from the ambient plasma. The first peak is the actual arc plasma wave front passing the probe. The significance of the second peak is not known at this time. This type of wave structure was typical of all data recorded during these tests.
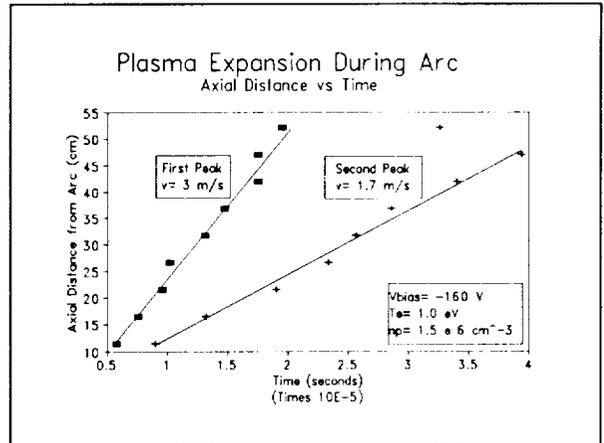
Figure 6   Arc Plasma Expansion Rate

The time for the arc plasma wave front to reach the Langmuir probe was plotted as a function of axial position and is shown in Fig. 6. This figure provides an indication of the velocity or the rate of expansion of the arc plasma into the ambient plasma. The electrons in the plasma which are more mobile than the ions are controlled by the electrostatic interaction of the ions. The velocities shown in Fig. 6 were computed by calculating the best linear fit to each set of data. The calculated velocities are an order of magnitude higher than ion acoustic velocities at these plasma conditions.

The average kinetic energy of an ion in the arc plasma can be calculated assuming every ion has the velocity shown in Fig. 6. Using the velocity of the first peak to calculate the kinetic energy in the first wave front assuming aluminum ions, an energy of 127 eV is obtained. The calculated kinetic energy is close to the energy one would expect a particle would obtain from the plate bias (i.e. -160 V). The difference between the calculated kinetic energy and the plate bias is an indication of the energy required to vaporize and ionize the aluminum to produce a plasma.

The intense arc plasma as seen from the data expands at a rate given by the amount of energy provided to the arc plasma by the electrons. The arc plasma expands at a velocity sufficient to cover large surface areas of the anodized aluminum structure over a short period of time. The ability of the arc plasma to discharge the surface of the anodized aluminum structure was examined by placing two plates in a plasma (See Fig. 3). The plates were isolated from each other inside the chamber, but connected outside. When one plate arced, the arc plasma would discharge the other plate passing current between the two.

727

Typical current and voltage data recorded during the two plate experiment are shown in Fig. 7. When this data was
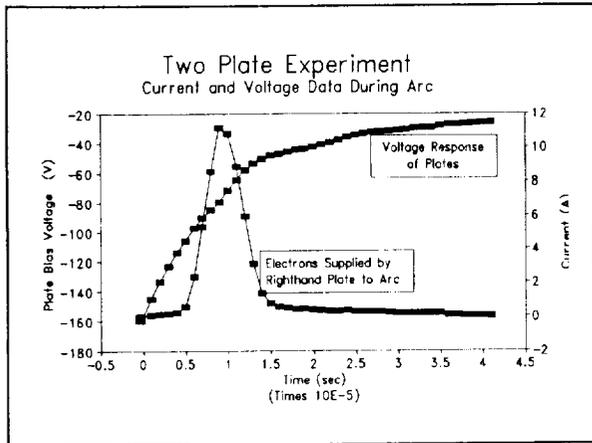
Figure 7   Two Plate Experiment Current and Voltage Data

acquired, an arc occurred on the lefthand plate shown in Fig. 3. The beginning of the voltage trace indicates the inception of the arc. In this figure the current is positive which was positive proof of electron flow from the righthand plate.

The time lag between the initiation of the arc and the peak in the current trace represents the time required for the arc plasma to be emitted from the lefthand plate and fully discharge the righthand plate. A simple calculation of the time for the plasma to travel the distance from the righthand plate to the left was performed. The two plates are 36 cm apart from the center of one plate to the center of the other. Because an arc could occur anywhere on the lefthand plate a range of distances must be considered. The range considered in this calculation was 32 cm to 40 cm, which is the distance from the near edge of the lefthand plate to the center of the righthand plate and the far edge to the center of each respective plate. Using the velocity indicated in Fig. 6, the time for the plasma to reach the center of the second plate ranged from 10 to 13 microseconds. This time is close to the lag time between the start of the arc and the peak in the current trace. This suggests that the arc plasma can expand and discharge the neighboring anodized aluminum plate.

The level electron currents being transferred between the two plates was also significant. The data presented in Fig. 8 is the current being supplied between the plates by either the right or lefthand plates, as indicated, after being discharged by the arc plasma. The plates, from which the data shown in this
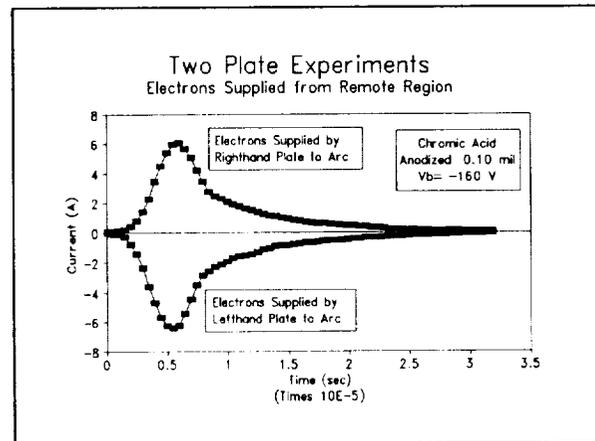
Figure 8   Current Traces from Either Plate

particular figure were measured, were both anodized aluminum with a 0.1 mil (2.54 microns) oxide layer. The current traces shown in Fig. 8 can be numerically integrated to calculate the charge transferred between the plates at a given voltage. From that information the experimentally measured capacitance can be computed. Also, the theoretical capacitance of the two plates can be calculated using flat plate theory assuming the aluminum as one conductor and the plasma as the other.

The numerical integration program used to compute the experimentally measured capacitance was calibrated by measuring the current response to several capacitors which were biased   -160 V and shorted to ground. Current and voltage data similar to that shown in Fig. 7 were recorded and numerically analyzed to calculate an experimentally measured capacitance. The capacitors experimentally measured were on the order of a few microFarads which is close to the capacitance of the plates in question. The experimentally measured capacitance agreed to within 10 % of the value listed on the label of the capacitors.

The results of the numerically calculated capacitance values obtained from the two plate experiments are compared with those calculated from flat plate capacitance theory in Table I. The capacitance data shown in Table I were computed for aluminum plates identically processed with two different anodic coating thicknesses which are indicated in the table. The experimental capacitance data are within 20 to 25 % of that calculated from theory. Two reasons for the difference in the two capacitance values are the uncertainty of the dielectric constant and the 10 % error already present in the integration program. The dielectric constant found in the literature was for aluminum-oxide not anodized aluminum. The

728

Comparison of Theoretical and Experimental Capacitance

Anodize Cond.	Exper. Cap. (uF)	Theory Cap. (uF)
0.05 mil Chromic Acid An.	0.49	0.68
0.10 mil Chromic Acid An.	0.27	0.38

Table I   Comparison of Theoretical and Experimental Plate Capacitance Values

Figure 9   Plasma Response to Simulated Arc

oxide coating produced during the anodization process does have some metallic impurities present. The results show the charge being transferred between the two plates is equal to that being stored in the plates. This suggests the arc plasma can completely discharge the stored charge in the surface once the two come in contact.

The expansion of the arc plasma can discharge a large area in a very short time. The size of the area that can be discharged is the present subject of an ongoing effort. However, it is apparent that physical limitations of spacecraft design (i.e. around corners or on the backside) will prevent the arc plasma from discharging an entire spacecraft. Therefore, experimental tests were performed to examine the interaction of the ambient plasma with the part of a spacecraft isolated from the arc plasma.

When an arc occurs on a spacecraft, the portion isolated from the arc plasma only experiences a positive voltage transient. Tests were conducted to measure the response of the ambient plasma to these simulated arcs by applying the same voltage transient to the anodized aluminum sample. This voltage transient was produced by biasing the sample 160 V negative of ground and immediately grounding the plate through a switch. The ability of the ambient electrons to discharge the surface of the plate was recorded as the amount of electron current flowing from the plate to ground. Because the ground test facility has a limited amount of plasma in the vacuum vessel, it was necessary to ground the hollow cathode plasma source in these tests to simulate the infinite source of electrons present in space. This is a viable approach because hollow cathode plasma sources are capable of several amperes of electron current.
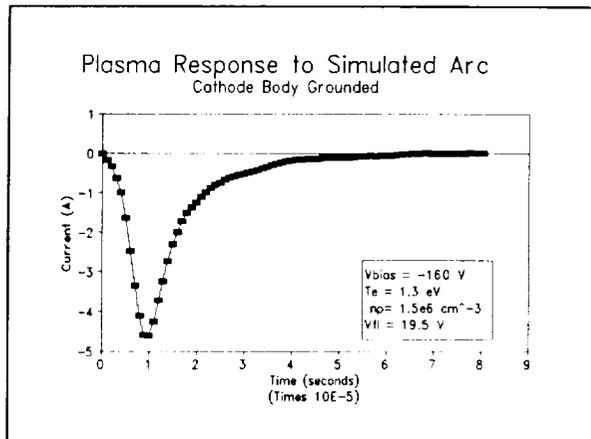
all data taken with the hollow cathode grounded. When the hollow cathode was not grounded, the current response to the simulated arc showed some unusual perturbations which suggested the plate was probably collecting all the electrons possible from the limited plasma volume with in the chamber. The peak current shown in this figure is on the order of amperes whereas the amount of thermal electron current the plate could collect is several orders of magnitude smaller (on the order of uA). In this transient case the ions are frozen and the Debye sheath which shields electric fields has not had time to develop. The time response on the current trace shown in this figure appears to be slower and the peak current smaller than data recorded when an arc occurred on a similar type of plate (see Fig. 7) without the hollow cathode grounded. This added time delay is due to the increased impedance caused by the absence of the conductive arc plasma. The data presented in this figure were also numerically integrated to determine the amount of charge measured during the simulated arc. This numerical calculation was then compared to that observed in the two plate experiments.

Figure 10 shows a comparison of the total integrated charge for both the two plate experiments and the simulated arc experiment. This figure is a plot of the total integrated charge obtained from numerical integration of current traces like those shown in Figs. 7, 8 and 9 versus the bias voltage on the plate minus the floating potential of the anodized surface. Because the hollow cathode was grounded during the simulated arc tests, the plasma floating potential was nearly 20 V. In the case of the two plate experiment the floating potential was about 2 V. The data

729
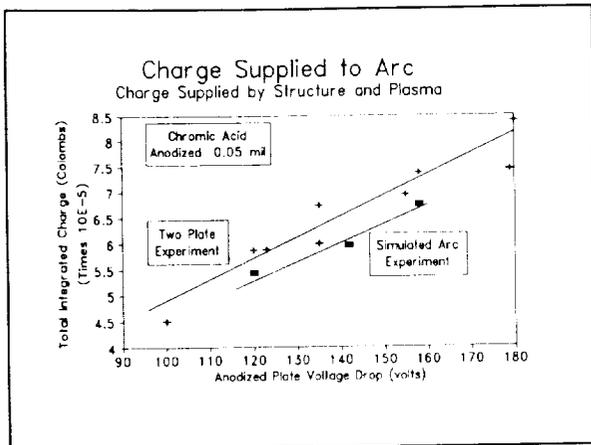
Figure 10   Comparison of Charge Supplied to Arc Between the Two Plate Experiment and the Simulated Arc Experiment

Figure 11   Peak   Currents   Caused   by Dielectric Breakdown

in this figure show the charge transferred during each experiment agree within experimental error. This result indicates that the ambient plasma is capable of discharging the entire anodized surface it comes in contact with in a very short period of time provided there is a sufficient supply of charged particles. Because space is an infinite well of charge particles, there should be ample supply to discharge any spacecraft in orbit.

SUMMARY

The requirement for higher powered spacecraft has produced new environmental interactions between the low earth orbit plasma and the structure. From the results of this work and that presented in [10], under the present design practice which grounds or references the structure to the negative side of the solar array, dielectric breakdown will occur on spacecraft whose potential relative to the plasma is greater than 80 V negative and the LEO plasma will be able to support large currents on the order of 100's or 1000's of amps depending on the capacitive capability of the spacecraft. Figure 11 is a reprint of the data shown in [10] to provide the peak current data versus capacitance that has been measured in the laboratory.

The electrons collected by the solar array from the ambient plasma build up in a capacitive nature in the skin of the structure. During the breakdown of the dielectric coating on the structure, there is sufficient energy to vaporize and ionize the aluminum structure at the arc site, producing a dense arc plasma. The arc plasma expands from the arc site out in every direction at a rate much greater than the ion acoustic velocity. As the arc plasma expands along the structure it neutralizes the ions which have collected

along the structure due to its capacitance. The amount of the structure that the arc plasma can discharge is determined by how far the arc plasma can travel in 10's, possibly 100's, of microseconds and have the needed current capacity. The exact distance is not known yet and is a subject of an ongoing effort.

The portion of the spacecraft which does not interact directly with the arc plasma will be discharged by the ambient plasma. The voltage transient the spacecraft experiences during the arc is sufficient to attract enough ambient electrons to discharge the entire stored charge in the structure. The arc will have to couple to the ambient plasma to provide a closed current path. The electron current attracted by the voltage transient from the ambient plasma is much greater than thermal electron currents.

REFERENCES

1.   MSFC Electrical Division Personnel, Private Communication.

2.   K.L. Kennerud, "High Voltage Array Experiments," NASA CR-121280, 1974.

3.   N.J. Stevens, "Investigation of High Voltage Spacecraft System Interactions With Plasma Environments," AIAA Paper No. 78-672, 1978.

4.   J.E. McCoy and A. Konradi, "Sheath Effects Observed on a 100 meter High Voltage Panel in Simulated Low Earth Orbit Plasmas," Spacecraft Charging Technology-1978, NASA CP-2071, 1978.

5.   N.T. Grier and N.J. Stevens, "Plasma Interaction Experiment (PIX) Flight Results," Spacecraft Charging Technology-1978, NASA CP-2071, 1978.

730

I

6. W. Kerslake and S. Domitz, "Neutralization Tests on the SERT II Spacecraft," AIAA Paper No. 79-2064, 1979.

7. N.T. Grier, "Plasma Interaction Experiment II (PIX-2): Laboratory and Flight Results," Spacecraft Environmental Technology-1983, NASA CP-2359, 1983.

8. N.T. Grier, "Experimental Results on Plasma Interactions with Large Surfaces at High Voltages," NASA TM-81423, 1980.

9. S. Jones, J. Stakus and D. Byers, "Preliminary Results of SERT II Spacecraft Potential Measurements Using Hot Wire Emissive Probes," AIAA Paper No. 70-1127, 1970.

10. M.R. Carruth, Jr., J.A. Vaughn, and P.A. Gray, "Electrical Breakdown of Space Station Freedom Surfaces," appears in the proceedings of the 26th Intersociety Energy Conversion Engineering Conference, August 4-9, 1991.

11. D.E. Siegfried, "A Phenomenological Model for Orificed Hollow Cathodes," NASA CR-168026, Dec. 1982.

Electrostatic Effects on Dust Particles in Space

by

Philip Leung

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

and

Ralph Wuerker

University of California
Los Angeles, California

The star scanner of the Magellan spacecraft experienced operational anomalies continuously during Magellan's journey to Venus. These anomalies were attributed to the presence of dust particles in the vicinity of the spacecraft. The dust particles, which were originated from the surface of thermal blankets, were liberated when the electrostatic force acting on them was of sufficient magnitude. In order to verify this hypothesis, an experimental program was initiated to study the mechanisms responsible for the release of dust particles from a spacecraft surface. In the experiments, dust particles were immersed in a plasma and/or subjected to ultra-violet irradiation. Results showed that the charging state of a dust particle was strongly dependent on the environment, and the charge on a dust particle was approximately 10^3 elementary charges. Consequently, in the space environment, electrostatic force could be the most dominant force acting on a dust particle.

Session E5: SPACE DEBRIS - I:
REVIEW, SPECIAL MODELS

Session Chair: F. Allahdadi

THE ASSESSMENT OF LONG-TERM ORBITAL DEBRIS MODELS

by

F. M. Jonas, PhD
K. W. Yates

ORION International Technologies, Inc.
300 San Mateo NE, Suite 200
Albuquerque, NM 87108

Abstract

This paper presents the results of on-going research being conducted for the U. S. Air Force Phillips Laboratory by ORION International Technologies, Inc. The purpose of this research is to assess existing long-term orbital debris models as a first step in the Air Force's effort to develop an Air Force long-term orbital debris modeling capability. Existing models are assessed against Air Force requirements for a long-term orbital debris model which can; (a) operate with the necessary accuracy at the relevant altitudes and orbital parameters, (b) benefit from new Air Force and non-Air Force debris measurements, and (c) accommodate current and future Air Force space scenarios. Model assessment results are shown for the NASA engineering model. The status of the NASA EVOLVE model assessment is discussed.

INTRODUCTION

This paper presents the results of studies conducted for the U. S. Air Force Phillips Laboratory, Albuquerque, NM, on long-term orbital debris models. The research was conducted by ORION International Technologies, Inc., as a task under contract F29601-89-C-0001. The objective of the research is to assess existing long-term orbital debris predictive models against Air Force (AF) needs for such a model which can: (1) operate with the necessary accuracy at relevant altitudes and orbital parameters; (2) benefit from new debris measurements; and (3) accommodate current and future AF space scenarios. Specifically, assessment results are shown for models developed by the National Aeronautics and Space Administration (NASA) Johnson Space Center (JSC). First, summary assessment results are presented for the NASA empirical engineering model developed for spacecraft in low earth orbit (LEO). The model formulation, sensitivities and uncertainties are discussed. Although the model does meet the NASA objective of designing a model that is easy to use by the design community, it does not meet the AF needs as stated above. Next, the status of the current investigation on the NASA EVOLVE computer code is summarized. EVOLVE is a more sophisticated empirical and physics based computational model designed to operate on both mini- and micro-computer systems. The EVOLVE model formulation and typical results will be shown.

BACKGROUND

The man-made orbital debris environment is becoming of increasing concern to the international space community. The majority of the man-made debris is in LEO, with the estimated total mass of orbital debris in LEO ranging anywhere from 1.5 to 3 million kg (Ref. 1, 2). Orbital debris consists of large objects such as non-functioning payloads to small objects such as paint flakes . Most of these objects are in high inclination orbits and pass one another at average relative velocities of 10 km/s (Ref. 2). The concern to space users ranges from catastrophic collision with large particles to surface erosion, especially of sensors, due to the smaller particles. Although the threat due to impact of orbital debris on space systems is not yet substantial, it is growing, and if left unchecked the future debris environment may indeed be extremely serious (Ref. 1,2). Currently, the best information on this environment comes from the Space Surveillance Network and returned spacecraft surfaces. It is estimated that the Space Surveillance Network currently cannot track objects less than 10 cm in diameter (Ref. 2). Only 5 percent of the approximate 6500 objects currently being tracked by USSPACECOM represent functioning operational spacecraft (Ref. 1, 2). 99 percent of the estimated orbital debris mass resides in these large trackable objects. The majority of the surface impact data to date comes from the Solar Maximum Mission Satellite (Solar Max) (Ref 3). Over 3 m^2 of surface area was returned for examination, recording 300 plus impacts during four plus years exposure at nominally 500 km and 28.5 deg orbital inclination. Orbital debris particles impacting these surfaces were estimated to have diameters ranging from 10^{-4} to 10^{-2} cm. Other limited data does exist giving some information on sizes around 1 cm; however, very little is known about this environment, other than at 500 km, and for sizes greater than 10 cm in diameter.

NASA ENGINEERING MODEL

NASA has developed an engineering model of the debris environment intended for spacecraft operating in LEO (Ref. 4 , 5). This empirically based model represents curve fits to available data. An assessment of this model by ORION shows that the model only be used at altitudes below 1000 km, and for debris sizes in the range $10^{-4} \leq d \leq 10^3$ cm. Orbital debris flux using this model is predicted using the following expression:

$$F(d,h,i,t,S) = H(d)\ \Phi(h,S)\ \Psi\ (i)\ [F_1(d)g_1(t) + F_2(d)g_2(t)] \tag{1}$$

Where,

$F(d,h,i,t,S)$ diameter d	=	Time averaged cumulative orbital debris flux (Impacts/m^2-yr) for debris particle sizes of and greater on a single sided randomly tumbling surface
d	=	Debris diameter (cm)
h	=	Orbital altitude (km)
i	=	Orbital inclination (deg)
t	=	Year (yr)
S the	=	13 month smoothed solar radio flux at 10.7 cm wavelength expressed in 10^4 Jansky (Jy) for previous year (t-1)
$H(d)$	=	$\sqrt{10^{exp}\left[-\ (\log_{10}d - 0.78)^2/(0.637)^2\right]}$
$\Phi(h,S)$	=	$\Phi_1(h,S)/[\Phi_1(h,S) + 1]$, Solar/Atmosphere effects
$\Phi_1(h,S)$	=	$10^{(h/200 - S/140-1.5)}$
$\Psi(i)$	=	Orbital debris inclination distribution (Fig. 1)
$F_1(d)$	=	$1.22\ (10^{-5})\ d^{-2.5}$, Small particle flux
$F_2(d)$	=	$8.1\ (10^{10})(d + 700)^{-6}$, Large particle flux
$g_1(t)$	=	$(1 + q)^{(t-1988)}$, Small particle growth where q = 0.02 (2%) until 2010 then q = 0.04 (4%)
$g_2(t)$	=	$1 + p(t-1988)$, Large particle growth where p = 0.05 (5%)
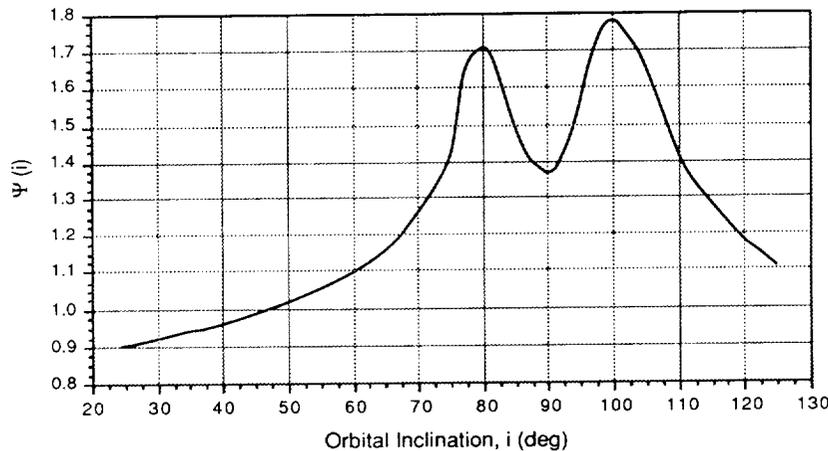
Figure 1. Orbital debris inclination distribution function, Y(i).

Predictions of the orbital debris environment using this model are shown in Figures 2 and 3 as compared to data. Note that above 1000 km the model predicts essentially constant flux (Fig. 2).
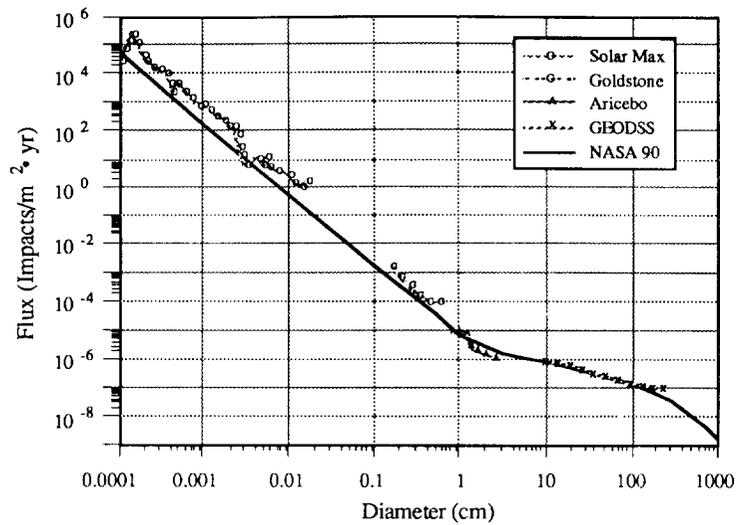
Figure 2. NASA 90 orbital-debris flux versus diameter, d (cm), compared with
data from Solar Max, Goldstone, Aricebo, and GEODSS
(h = 500 km, i = 28.5 deg, t = 1988, S = 140).
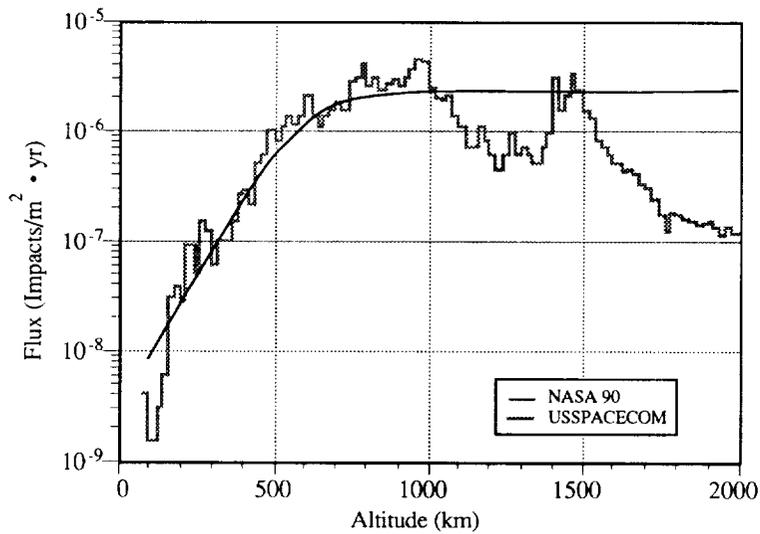
Figure 3. NASA 90 orbital-debris flux versus altitude, h (km), compared with
USSPACECOM data (d ≥ 10 cm, i = 60 deg, t = 1990, S = 200).

Engineering Model Sensitivity Study

A sensitivity study of the NASA engineering model was performed to quantify the model response to variation in the parameters d, h, i, t, S, p and q. The purpose of this study was to identify the most to least sensitive parameters as well as any anomalous or singular behavior in the model. This information is useful to the designer in knowing where small relative changes in the input can cause large variations in the output or vice-versa. In addition, if there is any singular behavior the user will know to avoid these regions.

The model sensitivities were determined by examining the first-order partial derivatives of F(d, h, i, t, S) with respect to each variable (which gives the relative change in the flux prediction with respect to changes in each variable). Because the majority of functions which make up F are smooth and continuous, it is expected that the first-order derivatives also will be smooth and continuous. Ψ (i) is based upon tabular values, but results in a smooth and continuous curve over the range 25 deg $\leq$ i $\leq$ 125 deg as shown in Figure 1. The exception is $g_1(t)$ where q changes value at the year 2010, however, $g_1(t)$ is smooth and continuous on either side of this discontinuity. Summary results of this study are shown in Figure 4 for the baseline values of h = 500 km, d $\geq$ 1 cm, i = 47 deg [Ψ(47 deg) = 1], t = 1995, S = 90, p = 5% and q = 2%. Each parameter was varied as follows:

- diameter $\partial F/\partial d$ vs. d   for $10^{-3} \leq d \leq 10^2$ cm
- altitude $\partial F/\partial h$ vs. h   for $100 \leq h \leq 2000$ km
- inclination $\partial F/\partial i$ vs. i   for $25 \leq i \leq 125$ deg
- time $\partial F/\partial t$ vs. t   for $1988 \leq t \leq 2010$
- solar activity $\partial F/\partial S$ vs. S   for $70 \leq S \leq 250$
- small particle growth $\partial F/\partial q$ vs. q for $0 \leq q \leq 0.20$
- large particle growth $\partial F/\partial p$ vs. p for $0 \leq p \leq 0.20$

Figure 4 plots the variation of each parameter as a function of the normalized parameter for comparison purposes.
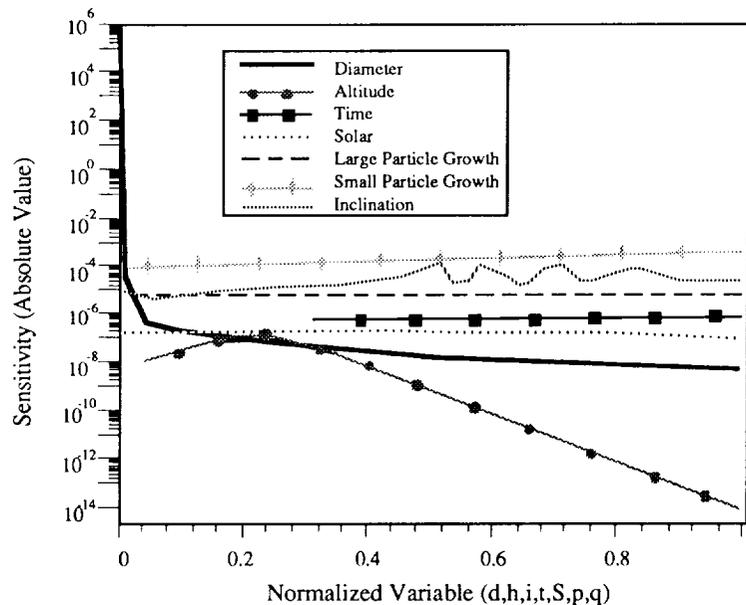
Figure 4. Flux Model Sensitivities, Summary Comparisons (Case II).
Baseline Values: h = 500 km, d = 1 cm, i = 47 deg, t = 1995, S = 90,
p = 5%, q = 2%

The rank order of model sensitivities and comments pertaining thereto are as follows:

Parameter	Relative Order of Magnitude/Comments
d (diameter)	~ 10^6 at d = 0.001 cm ~ 1 at d = 0.05 cm ~ 10^{-9} at d = 100 cm

737

Flux prediction becomes undefined as d → 0

q (small particle growth rate)	~ 10^{-4}, slightly increasing with increasing q
i (orbital inclination) (large particle growth rate)	~ 10^{-5} to 10^{-6}, reflects Ψ (i) variation
p (large particle growth rate)	~ 10^{-5} to 10^{-6} sensitivity is constant throughout range
t (time)	~ 10^{-7} to 10^{-8}
s (solar)	~ 10^{-7} to 10^{-8}
h (altitude)	~ 10^{-8} (h = 500 km) to 10^{-14} (h = 2000 km), predicts constant flux for h > 1000 km

Engineering Model Uncertainty Study

An uncertainty analysis of the NASA model was performed to quantify the accuracy of the model predictions. The study was conducted using a propagation of error analysis that examines how the root mean squared error in the data affects the results obtained from the use of the model. The uncertainty or error in the flux prediction, ε_F, was estimated using the following relationship:

$$\varepsilon_F = \pm \left\{ (\varepsilon_{FM})^2 + (\varepsilon_{FH})^2 + \left(\frac{\partial F}{\partial p} \varepsilon_p \right)^2 + \left(\frac{\partial F}{\partial q} \varepsilon_q \right)^2 + \left(\frac{\partial F}{\partial S} \varepsilon_S \right)^2 \right\}$$

(2)

where

ε_F	=	Uncertainty in F, e.g., F± ε_F
ε_{FM}	=	Uncertainty in F because of statistical and measurement uncertainties primarily as a function of debris diameter
ε_{FH}	=	Uncertainty in F with respect to altitude due in part to the difficulty of determining F for debris in highly elliptical orbits
ε_p	=	Uncertainty in large particle growth
ε_q	=	Uncertainty in small particle growth
ε_S	=	Uncertainty in the level of solar activity

Table 1. NASA Estimated orbital-debris uncertainties 90% confidence (Ref. 5).

Parameter		Estimated Uncertainty
Flux Measurements	d ≥ 10 cm	1.5 to 0.5 x Flux
	0.05 < d < 10 cm	3.0 to 0.33 x Flux
	d ≤ 0.05 cm	2.0 to 0.5 x Flux
Altitude Distribution,	d ≥ 10 cm	2.0 to 0.5 x Flux
	1 < d < 10 cm	3.0 to 0.33 x Flux
		1.5 to 0.5 x Flux for 300 < h < 700 km
	d ≤ 1 cm	5.0 to 0.2 x Flux per every 200 km away from h = 500 km
Orbital-debris Growth, q		4% to 10%
		0% to 20%
Solar Activity, S		Use max and min values for nominal solar cycle where typically 70 ≤ S ≤ 210

738

I
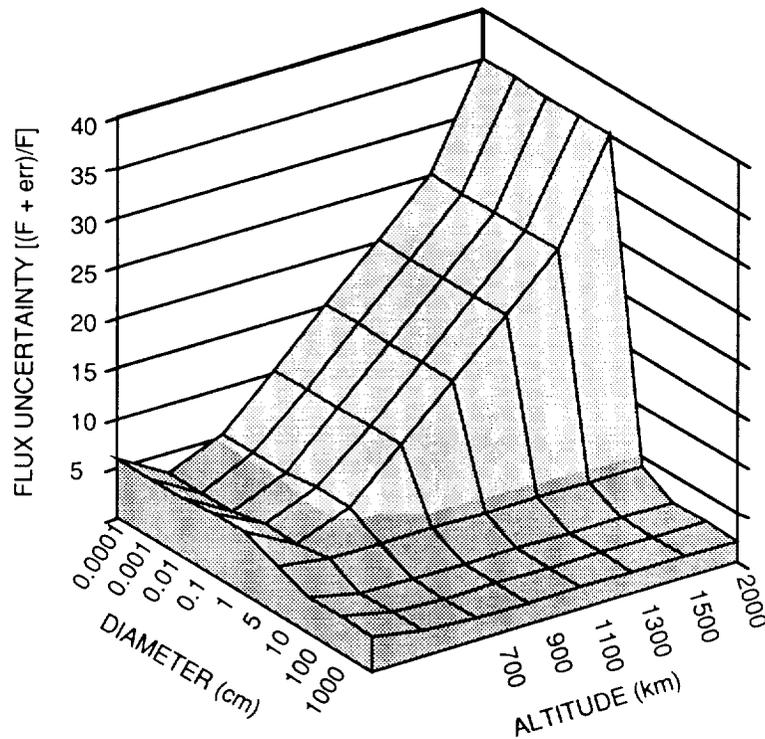
Figure 5. Flux model uncertainty vs. altitude and diameter (i = 30 deg, t = 1995,
S = 90)

Figure 5 shows the summary results, [(F + ϵ_F)/F], versus altitude and diameter at an orbital inclination of 30 deg for the year 1995. As can be seen in this figure, the best model predictions (factor of 2-4) occur at 500 km and/or for particle sizes greater than 5 cm in diameter. The uncertainty for the smaller particles and higher altitudes is seen to be at least an order of magnitude. Similar results are obtained for all orbital inclinations. As a result of the sensitivity and uncertainty analyses it is recommended that use of the model be restricted to diameters, 10^{-4} cm ≤ d ≤ 10^3 cm, and altitudes, h ≤ 1000 km, through the year 2010.

NASA EVOLVE (Evolutionary) MODEL

The NASA EVOLVE model (Ref. 6) is a NASA in-house numerical model developed to predict the man-made orbital debris environment. A schematic of the EVOLVE model is shown in Figure 6. The EVOLVE model uses USSPACECOM data (intact objects) to generate a historical data base covering the years 1957 - 1989. To project future traffic, the program uses data files containing anticipated launches determined from mission and traffic models. Any combination of mission models may be employed to project traffic for U.S. civil, DoD, Soviet, and other foreign agencies. Currently, the program uses a data base for 1990 - 2010 comprised of the U.S. Civil Needs Data Base (U.S. civilian traffic including shuttle missions) and a copy of the 1988 historical data to simulate DoD and foreign traffic. Historically documented breakups are included in the historical data base from 1957 - 1989. Explosions and/or collisions are determined stochastically and added to the deterministic flux environment. Objects are allowed to decay as a function of their interaction with the atmosphere. The program ultimately predicts the flux environment as a function of altitude and particle size for any given time. Sample results obtained by ORION using EVOLVE are shown in Figure 7.

739
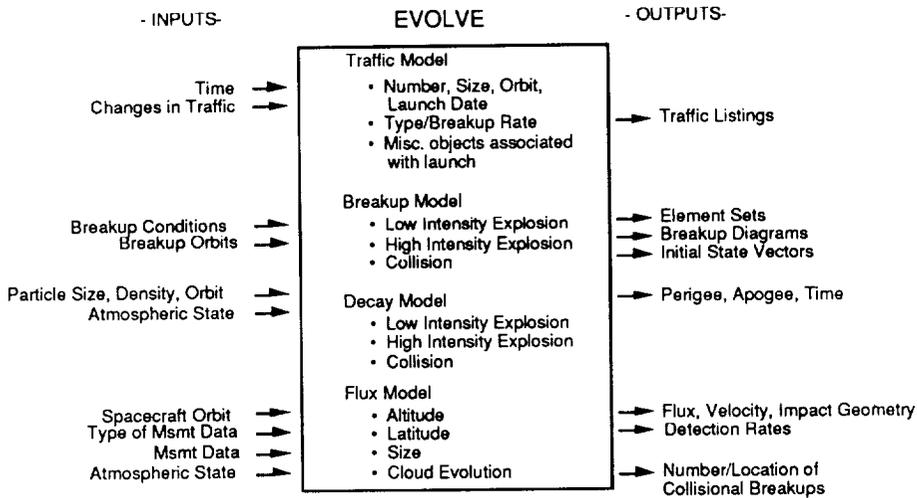
Figure 6. EVOLVE Schematic
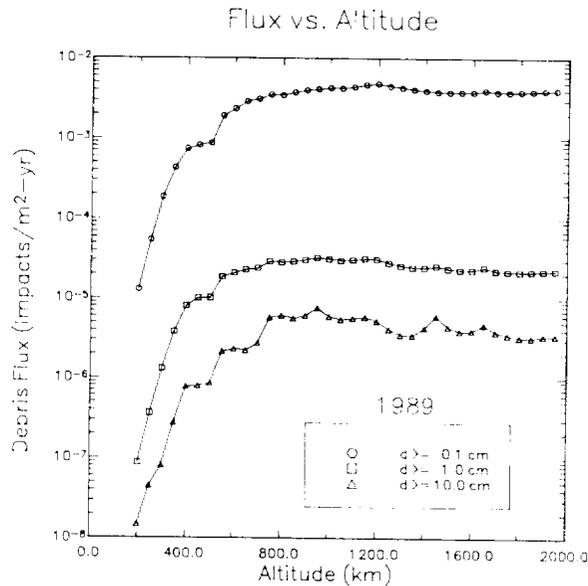
Figure 7. Debris Flux vs. Altitude (EVOLVE), 1989.

EVOLVE Assessment

The assessment of the NASA EVOLVE model is just beginning. It is planned to first dissect the EVOLVE code into its constituent parts and examine the methodologies used to calculate the various effects, and identify the critical parameters and assumptions. Next, a sensitivity study will be performed. In this analysis the code will be examined at different levels. At the upper most level the flux sensitivity to altitude, diameter and time will be examined. The sensitivity of the results to the output of various subroutines will be studied next, and finally the sensitivity of the flux prediction to critical parameters within those subroutines. The final part of the assessment will quantify the model uncertainties as compared to data of this environment. The model means and variances will be established. The assessment is to be conducted through December of this year with the results published in a Phillips Laboratory technical report early in 1992.

CONCLUSION

Existing long-term orbital debris models are being assessed against Air Force needs for a model that can (a) operate with high accuracy at the relevant altitudes and orbital parameters, (b) benefit from new Air Force and non-Air Force debris measurements, and (c) accommodate current and future Air Force Space scenarios. The models being assessed are the NASA engineering model and the NASA EVOLVE model. The assessment of the NASA engineering model has been completed. It has been concluded that this model does not meet the Air Force requirements. It should be noted however that the NASA engineering model does meet the NASA goal to design a model that is easy to use by engineers. It also has helped to establish threat levels for the first time and is pointing the way on where and how knowledge of the orbital debris environment needs to be improved. The assessment of the NASA EVOLVE model is just beginning. The results of this latter assessment will be available this fall (1991) with the final report available in 1992.

REFERENCES

1. Space Debris, A Report from the ESA Space Debris Working Group, European Space Agency, 1988.
2. Report on Orbital Debris, by Interagency Group (Space), for the National Security Council, Washington D.C., February 1989.
3. Laurance, M.R., and Brownlee, D.E., "The Flux of Meteoroids and Orbital Space Debris Striking Satellites in Low Earth Orbit," Nature, Vol. 323, 11 September 1986.
4. Kessler, Donald J., Reynolds, Robert C., and Anz-Meador, Phillip D., Orbital Debris Environment for Spacecraft Designed to Operate in Low Earth Orbit, NASA TM 100471, April 1989.
5. Proposed Revision to Section 8, "Meteoroids and Orbital Debris," Space Station Program, Natural Environment Definition for Design, NASA SSP 30425, September 1990.
6. Reynolds, Robert C., Documentation of Program EVOLVE: A Numerical Model to Compute Projections of the Man-made Orbital Debris Environment, System Planning Corp., SPC Report No. 0D91-002-U-CSP, February 1991.

SPACE DEBRIS CHARACTERIZATION

IN SUPPORT OF A SATELLITE BREAKUP MODEL

Bryan H. Fortson, James E. Winter,
and Firooz A. Allahdadi
Space Kinetic Impact and Debris Branch, Phillips Laboratory
Kirtland AFB, New Mexico 87117-6008

ABSTRACT

The Space Kinetic Impact and Debris Branch has begun an ambitious program to construct a fully analytical model of the breakup of a satellite under hypervelocity impact. In order to provide empirical data with which to substantiate the model, debris from hypervelocity experiments conducted in a controlled laboratory environment has been characterized to provide information on its mass, velocity, and ballistic coefficient distributions. Data on the debris has been collected in one master data file, and a simple FORTRAN program allows users to describe the debris from any subset of these experiments that may be of interest to them. A statistical analysis has been performed, allowing users to determine the precision of the velocity measurements for the data. Attempts are being made to include and correlate other laboratory data, as well as those data obtained from the explosion or collision of spacecraft in low earth orbit.

INTRODUCTION

Characterization of debris from hypervelocity impact events is an important prerequisite for analytical modelling of those events. One feature of a useful model would be its ability to predict the characteristics of the debris cloud produced by the impact, and this feature cannot be evaluated without a body of empirical data with which to compare the analytical predictions. The Space Kinetic Impact and Debris Branch (SKID) has conducted a debris characterization program, using hypervelocity impact debris produced in antisatellite experiments. Our approach includes some novel ideas from which the orbital debris community may benefit.

WEIGHING AND CATEGORIZING OF DEBRIS

SKID has characterized debris from hypervelocity shots conducted at the Naval Research Laboratory (NRL) and at the University of Dayton Research Institute (UDR-I). Most debris comes in plastic bags of two sizes, large and small. The large bags contain debris that has been swept from the floor of the impact chamber, and include a considerable amount of extraneous material. The smaller bags contain debris recovered from the inside of the impact target, and do not contain as much unwanted material.

The characterization begins with the sorting out of the metallic debris, from the non-metallic debris; this is done by hand. This provides another opportunity to remove extraneous material. The metallic debris is then weighed on the microbalance at Phillips Laboratory's metrology facility. This microbalance has an accuracy of 0.004 g. Debris particles with mass less than 10 times the accuracy of the scale are not weighed. The mass of the particle, as well as all pertinent information about the shot from which it came (shot number, impact angle, nominal projectile velocity, projectile material), is recorded in a master data file.

Some debris was collected in catcher material during the test, for the purpose of measuring its velocity. This debris is removed from the catcher material with tweezers, and the distance that it travelled into the catcher material is measured. This distance is used to determine the velocity at which the particle had travelled, using methods to be discussed later in this paper. The mass, velocity, and shot information are then included in the master data file.

CALCULATION OF DISTRIBUTION CURVES

Calculation of mass distribution curves is performed by the FORTRAN-77 code DEBRIS.F. The program selects the data of interest to the user, and then generates a mass distribution plot, which is written to a file called DISTRIB.DAT. The data can be plotted in two ways. In the first approach, the dependent variable is mass, and the independent variable is the total number of debris particles larger than that mass. Such a plotting method is often used in the space debris community, for example, in References 1 and 2. In the second approach, the dependent variable is still mass, but the particles are separated into "bins." For example, a bin might contain all particles with mass between 0.1 g and 0.3 g. If 12 particles are found to lie in this range, the resulting data point would be (0.2, 12). This method of plotting is similar to the approach used by the debris hazard modelling computer program IMPACT[3].

DEBRIS.F also fits a straight line and a parabola to the curve, and performs an F-test (as described in Reference 4) to determine if the higher-order fit is needed. The least-squares coefficients and the F-test results are written to a file called FITS.DAT. The source code can also generate velocity and ballistic coefficient distributions. DEBRIS.F, and the master data file DEBDAT, are installed on the CRAY-2 supercomputer at Kirtland AFB, whence they can be easily downloaded onto $5\frac{1}{4}$" floppy disks.

It has been noted by McKnight[5] that a straight-line fit to a mass distribution does not make physical sense: if the line is extrapolated to the left, it is seen to imply the presence of an infinite number of infinitesimal particles. It makes more sense for the curve to have a "knee." Such a "knee" can be introduced with a bilinear least-squares fit, which has been derived as a part of this effort, and is available along with the linear and parabolic fits as part of the DEBRIS.F code.

Consider the situation shown in Figure 1: a bent straight line is to be fit to a set of data. The line is defined by four parameters: the intercept β_0, the two slopes β_1 and β_2, and the x-coordinate of the break point, x_0. Consider the case in which x_0 is given. A least-squares fit is obtained by minimizing the residual sum of squares SS(res), given by
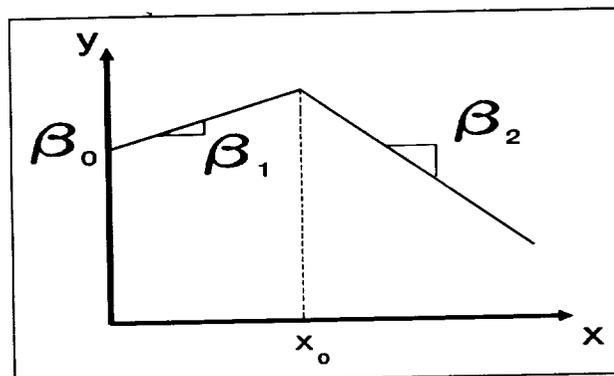
Figure 1. Definition of bilinear coefficients

$$SS(res) = \sum_{i=1}^{n_s} [y_i - (\beta_0 + \beta_1 x_i)]^2$$

$$+ \sum_{i=n_s+1}^{n_T} [(y_i - \beta_0 - \beta_1 x_0) - \beta_2 (x_i - x_0)]^2 \quad (1)$$

where $x < x_0$ for data points 1 through n_s, and n_T is the total number of data points. This is minimized by setting the partial derivatives of SS(res) with respect to the parameters equal to zero:

$$\frac{\partial SS(res)}{\partial \beta_0} = \frac{\partial SS(res)}{\partial \beta_1}$$

$$= \frac{\partial SS(res)}{\partial \beta_2} = 0 \quad (2)$$

Substituting equation 1 into equations 2 leads to a system of three linear equations in three unknowns:

$$[\alpha] \begin{Bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{Bmatrix} = [\gamma] \quad (3)$$

where $[\alpha]$ is a symmetric 3×3 matrix:

$$\alpha_{11} = n_T \quad (4)$$

$$\alpha_{12} = n_N x_0 + \sum_{i=1}^{n_s} x_i \quad (5)$$

743

$$\alpha_{13} = -n_N x_0 + \sum_{i=n_s+1}^{n_T} x_i \qquad (6)$$

$$\alpha_{22} = n_N x_0^2 + \sum_{i=1}^{n_s} x_i^2 \qquad (7)$$

$$\alpha_{23} = -n_N x_0^2 + x_0 \sum_{i=n_s+1}^{n_T} x_i \qquad (8)$$

$$\alpha_{33} = \sum_{i=n_s+1}^{n_T} (x_i - x_0)^2 \qquad (9)$$

the vector $[\gamma]$ is defined as

$$[\gamma] = \left\{ \begin{array}{c} \sum_{i=1}^{n_T} y_i \\ \sum_{i=1}^{n_s} x_i y_i + x_0 \sum_{i=n_s+1}^{n_T} y_i \\ \sum_{i=n_s+1}^{n_T} (x_i - x_0) y_i \end{array} \right\} \qquad (10)$$

and n_N is the number of data points for which $x \geq x_0$. This system can now be solved for β_0, β_1, and β_2. Note that this solution requires a "guess" for x_0, but not for the y-value of the break point. The FORTRAN code DEBRIS.F performs 1001 bent-line fits to a given data set, using 1001 evenly-spaced guesses for x_0. The fit with the best correlation coefficient is retained.

This bilinear fit is particularly useful for applying the bilinear exponential model of Grady and Kipp[6]. Reference 6 proposes an equation of the form

$$y = A_1 e^{-B_1 x} + A_2 e^{-B_2 x} \qquad (11)$$

This equation cannot be fit to a set of data using simple least-squares means; a more sophisticated, iterative approach is required. However, equation 11 approximates the bent-line fit if the x-axis is a linear scale, and the y-axis is a logarithmic scale. Thus, the bent-line fit can be applied to the model of Reference 6.

Figure 2 shows a mass distribution from DEBRIS.F, with the linear and parabolic fits employed. Data from all shots that were characterized was combined into one file, which is plotted here. The x-coordinate is a chosen mass M. The y-coordinate is the number of particles observed to have a mass greater than M. The F-test performed by the program indicates that
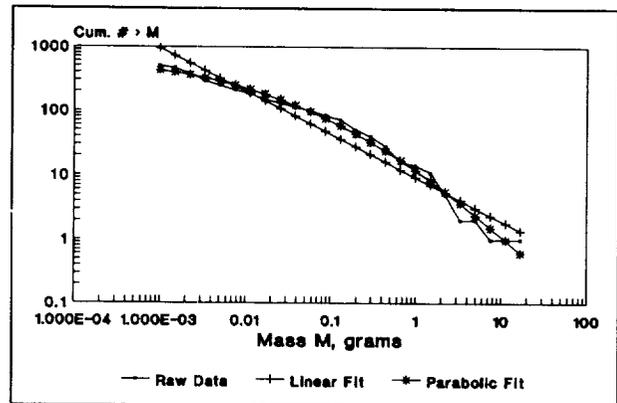
Figure 2. Mass distribution curve.

the accuracy gained by using the parabolic fit, as opposed to the linear fit, is statistically significant in this case.

Figure 3 shows the same data, separated into bins of uniform width on a logarithmic scale. (Bins can also be of uniform width on a linear scale.) The x-coordinate of each data point corresponds to the center of the bin. Linear and parabolic fits are also available with this type of plot.
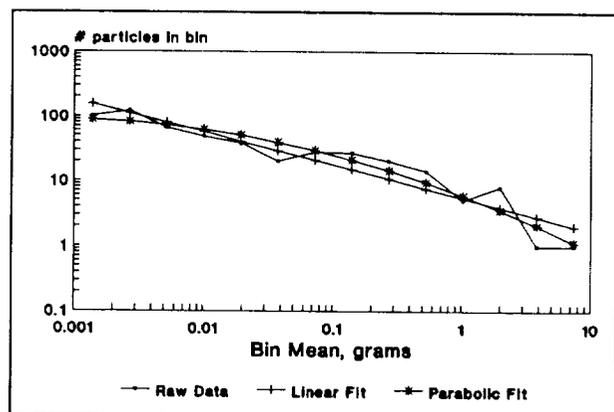
Figure 3. Sample plot with data separated into bins

Figure 4 shows the same data, plotted with the bilinear fit routine. At first glance, the fit does not appear to be very good. However, a closer look shows that the linear scale on the x-axis leads to a

concentration of data points near the left-hand side of the graph. Once these points are recognized, the fit is seen to be quite good.

Figure 4 also shows that, when the distribution data is plotted with a linear x-axis, the resulting curve is almost vertical at low values of x. For some test data, this steepness caused numerical stability problems in the bilinear fit routine, which resulted in negative values of R^2 for some or all of the trial fits. These stability problems were decreased, but not eliminated, by restricting the range of the trial values of x_0. The problem was solved by implementing a completely pivoted Gaussian elimination routine to solve equation 3. DEBRIS.F had previously solved equation 3 with an unpivoted Gaussian elimination.

Figure 5. Mass distribution plots for individual tests.
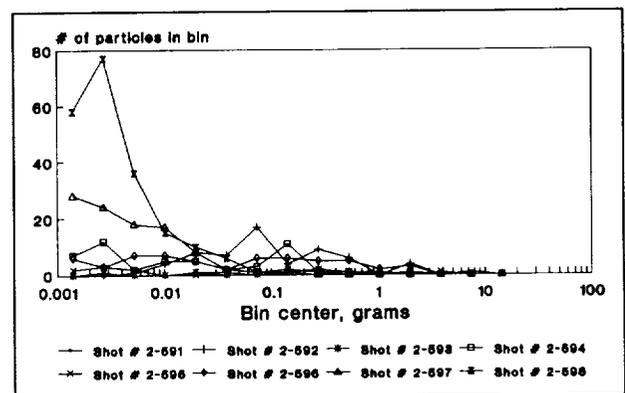
Figure 6. Individual mass distribution plots, using bin format.

Figure 4. Mass distribution plot with bilinear fit.

Figures 5 and 6 show individual plots for all tests, using the plotting formats of Figures 2 and 3. Figure 6 used a linear vertical axis, due to the presence of some zero values in the data. This format also emphasizes the differences in the data from the various tests. The most interesting result is the considerable variation among the curves shown in these two figures. These shots are actually quite similar; yet, their debris distributions are quite different. It should be noted that, while no debris particles of mass less than 0.04 g were weighed, this limitation would not change the values given in Figures 5 and 6; it would merely allow the curves to be extended to the left. Thus, the variability shown in these figures is real.

The flexibility of this approach would also be useful in studies such as Reference 2. McKnight and Brechin determined the linear fit coefficients for the mass distributions of several impact events, and fit a curve to these coefficients in an effort to create a master equation to describe the mass distributions of a wide range of impact events. Using the SKID method of cataloguing the debris, a large body of data could be built up and easily analyzed. Effects of varying parameters (such as projectile mass, projectile velocity, projectile material, etcetera) on the linear or parabolic fit parameters could be observed.

STATISTICAL ANALYSIS OF VELOCITY MEASUREMENTS

Velocity measurements are taken using the results of Malick's empirical study[7]. Malick fired test projectiles of four different materials at wallboard, using a

745

range of known velocities, masses, impact
angles (see Figure 7), and presented par-
ticle areas.

Malick then fit a curve to his data. The
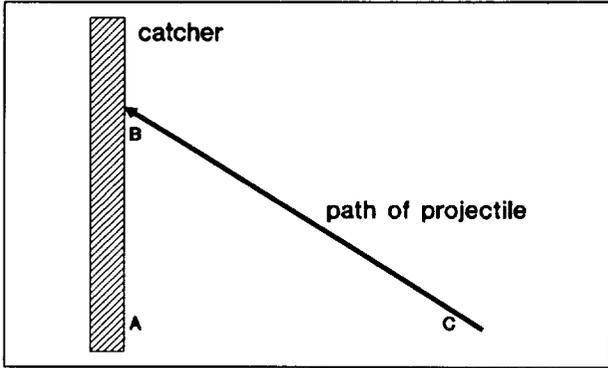
Figure 7. Definition of impact angle θ
($\triangle ABC$)·

resulting equation was:

$$V = \frac{1537 \; f_v \; t^{0.8091}}{m^{0.3336} (\cos \theta)^{0.5419}} \qquad \textbf{(12)}$$

where V is the particle velocity in ft/-
sec, f_v is a function of the projectile
density, t is the thickness of particle
board penetrated, in inches, m is the
weight of the particle, in grains, and θ
is the impact angle. The multiple corre-
lation coefficient for the fit of this
curve to the data is 0.88. Equation (12)
applies only to the case of approximately
cylindrical projectiles, for which

$$A \approx cm^{2/3} \qquad \textbf{(13)}$$

where A is the average presented area of
the particle, and c is a constant.

Figure 8 shows the results of fitting a
power-law curve to Malick's data for f_v as
a function of density. The resulting
curve is:
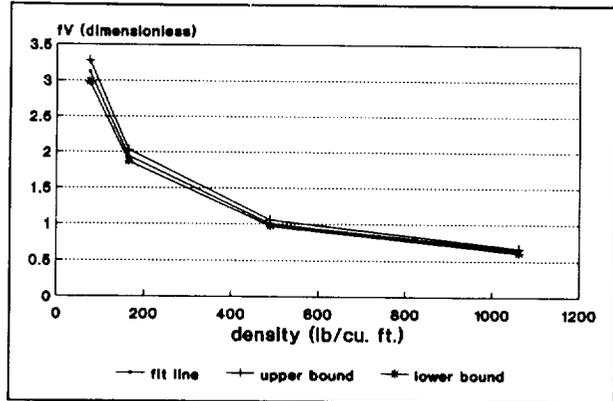
Figure 8. Power-law fit to Malick's data

$$f_v = \frac{41.2}{\rho^{0.6}} \qquad \textbf{(13)}$$

where ρ is the projectile density in lb/-
ft^3. This curve fits the common logarithms
of Malick's data with a correlation coef-
ficient of $R^2=0.9996$. Figure 8 shows a 95%
prediction confidence interval drawn about
this curve, illustrating the reliability
of interpolations of f_v made on the basis
of it. Incorporating equation (13) into
equation (12) gives

$$V = \frac{6.332 \times 10^4 \; t^{0.8091}}{m^{0.3336} \rho^{0.6} (\cos \theta)^{0.5419}} \qquad \textbf{(14)}$$

It is important to know the reliability of
velocity measurements made with Malick's
equation. A prediction confidence inter-
val for these measurements can be con-
structed, using the approach found in
Bancroft and Han.[8]

A multiple linear regression used to ob-
tain equation (14) would begin with an
equation of the form

$$\log V'_i = b_0 + b_1 \log \rho_i$$
$$+ b_2 \log t_i + b_3 \log m_i \qquad \textbf{(15)}$$
$$+ b_4 \log (\cos \theta_i)$$

where V'_i is the least-squares approxima-
tion to the actual velocity V_i. For conve-
nience, rewrite equation (15) as

$$\log V'_i = b_0 + b_1 x_{1i} + b_2 x_{2i}$$
$$+ b_3 x_{3i} + b_4 x_{4i} \qquad \textbf{(16)}$$

746

where the following abbreviations are used:

$$\log \rho = x_1 \quad (17)$$

$$\log t = x_2 \quad (18)$$

$$\log m = x_3 \quad (19)$$

$$\log (\cos \theta) = x_4 \quad (20)$$

Now, define the matrix [C] to be the inverse of the matrix [D], for which

$$D_{kl} = \sum_{i=1}^{n} x_{(k-1)i} x_{(l-1)i} \quad (21)$$

where all summations are taken over all n data points $\{V_i, x_{1i}, x_{2i}, x_{3i}, x_{4i}\}$, and x_0 is defined as 1. Furthermore, define the quantity S^2:

$$S^2 = \frac{1}{n-p} \sum_{i=1}^{n} (\log V_i - \log V'_i)^2 \quad (22)$$

where p is the number of degrees of freedom in the model. For this case, p=5; there are four coefficients in the regression equation, plus one constant. S^2 is an unbiased estimator of the variance of the errors of the data relative to the least-squares fit, assuming those errors to be normally distributed. Under this assumption, a $100(1-\alpha)\%$ confidence interval for V_i is given by

$$V'_i \pm t_{(\alpha/2, n-p)} S\sqrt{\{X\}^T [C] \{X\}} \quad (23)$$

where

$$\{X\} = \{ 1 \quad x_1 \quad x_2 \quad x_3 \quad x_4 \}^T \quad (24)$$

and $t_{(\alpha/2, n-p)}$ is the $100(1-\alpha/2)\%$ point of Student's t-distribution, with n-p degrees of freedom. Tables of this distribution are readily available in most probability texts. A $100(1-\alpha)\%$ prediction interval is given by

$$V'_i \pm t_{(\alpha/2, n-p)} S\sqrt{1 + \{X\}^T [C] \{X\}} \quad (25)$$

Thus, it is necessary to calculate the estimator S^2 and the matrix [C] from Malick's original data, after which the desired confidence intervals can be calculated by plugging the vector {X} into the appropriate equation.

Converting equation (14) to the MKS system gives

where V is now measured in m/sec, ρ is in kg/m^3, t is in meters, and m is in kilo-

$$V = \frac{1.108 \times 10^3 \ t^{0.8091}}{m^{0.3336} \rho^{0.6} (\cos \theta)^{0.5419}} \quad (26)$$

grams. With all data available from Reference 7 converted to MKS, the matrix [C] was calculated, and, using equation (26) as the fit to the data, the value of S was determined to be

$$S = 0.144 \quad (27)$$

Two facts about the data of Reference 7 are worth noting. First, the majority of shots were fired at impact angles of zero degrees, leaving relatively little information on the variation of V with θ. This is reflected in the large value of the coefficient C_{55} in the matrix [C], which indicates that the width of the confidence interval is very sensitive to θ. Second, the highest impact velocity recorded for valid data is about 9000 ft/sec, or about 2.7 km/sec. Malick performed some shots at higher speeds, but considered the resulting data invalid because of excessive breakup or deformation of the projectiles. Thus, the average velocity for Malick's shots is probably 1.0 - 1.5 km/sec. Figure 9, a schematic of the confidence and prediction bands about a simple least-squares straight-line fit, shows why this is important. The confidence band, which is described by equation 23, is tightest at the mean x-value of the data. Its width varies as the square of the distance from this point along the x-axis. Because of the linear nature of the fit, one can make a similar statement regarding the variation with distance from the mean y-value of the data. For the same confidence level, the prediction band, described by equation 25, behaves in a similar manner, but is much wider. The geometrical interpretation is more complex for a multiple regression. However, one can say that Malick's fit is probably most accurate for velocities of approxi-

mately 1.0 - 1.5 km/sec, and that the ac-
curacy of this method decays for values
significantly outside of this range.
Thus, it is important to have a quanti-
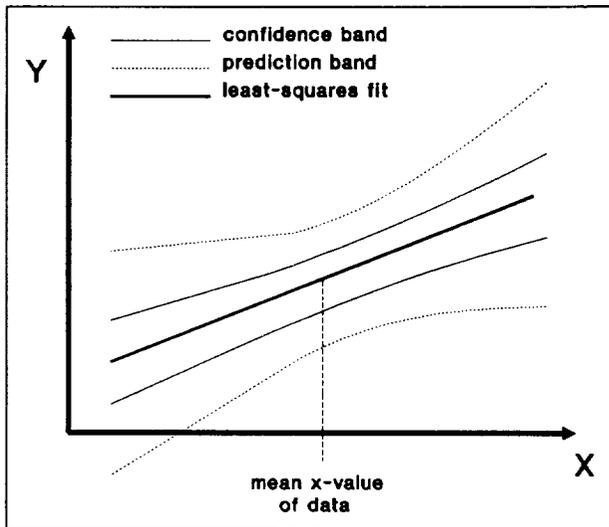
Figure 7. Schematic of confidence and
prediction intervals.

tative measure of the accuracy of the
velocity values thus obtained.

As an example, consider the first data
point given in Reference 7. The velocity
is seen to be 966 ft/sec, or 294 m/sec.
Converting the mass, density, and thick-
ness to the MKS system, and using equation
(25), Malick's curve fit is seen to indi-
cate a velocity of 493 m/sec for this
projectile. This illustrates a difficulty
resulting from fitting a curve to the
logarithm of the velocity data: Malick's
velocity estimate is in error by almost
70%, but this results from a creditable 9%
error in estimating the logarithm of the
velocity.

Using the definitions of equations (17) -
(20), and substituting into equation (23),
a 90% confidence interval for this data
point is found to be

$$442 m/\sec \leq V \leq 542 m/\sec \qquad \textbf{(28)}$$

However, a 90% prediction interval for
this same projectile, calculated using
equation (25), would be much larger:

$$281 m/\sec \leq V \leq 853 m/\sec \qquad \textbf{(29)}$$

CONCLUSION

Since the appropriate statistical parame-
ters have been calculated, one can now use
equations (23) and (25) to determine the
accuracy of any velocity prediction made
using Malick's method.

The method of creating debris character-
ization plots is very versatile. The data
file collected for this program can be
added to at any time, and the computer
code that reads it is simple and easily
modified. This flexible approach will
allow this database to be expanded and
restudied as new issues arise in debris
characterization.

ACKNOWLEDGEMENTS

REFERENCES

1. Maclay, T., and Ilinga, M., "Analysis
of Shot CU-5272 Fragments," COLORADO CEN-
TER FOR ASTRODYNAMICS RESEARCH IN TERNAL
RESEARCH REPORT NO. 153-6416, University
of Colorado at Boulder, June, 1990.

2. McKnight, Darren, and Brechin, Chris-
topher, "Debris Creation Via Hypervelocity
Impact," PAPER NO. AIAA-90-0084, American
Institute of Aeronautics and Astronautics,
Washington, DC, 1990.

3. Sorge, M.E., "Space Debris Hazard
Software: Program IMPACT Version 1.0 B,"

AEROSPACE REPORT NO. TOR-0091(6909-04)-1,
The Aerospace Corporation, El Segundo, CA,
January, 1991.

4. Miller, Irwin, and Freund, John E.,
PROBABILITY AND STATISTICS FOR ENGINEERS,
Prentice-Hall, Englewood Cliffs, New Jer-
sey, 1977. Pp. 319-320.

5. McKnight, Darren, private communica-
tion, January 1991.

6. Grady, D.E., and Kipp, M.E., "Geomet-
ric Statistics and Dynamic Fragmentation,"
JOURNAL OF APPLIED PHYSICS, vol. 58, no.
3, pp. 1210-1222.

7. Malick, D., "The Calibration of Wallboard for the Determination of Particle Speed," Ballistic Analysis Laboratory TECHNICAL REPORT NO. 61, May, 1966, The Johns Hopkins University.

8. Bancroft, T.A., and Han, Chien-Pai, STATISTICAL THEORY AND INFERENCE IN RESEARCH, Marcel Dekker, Inc., New York, New York, 1981. Pp. 277-289.

ANALYSIS OF ENERGY DISSIPATION AND DEPOSITION IN ELASTIC BODIES IMPACTING AT HYPERVELOCITIES

David F. Medina and Firooz A. Allahdadi

Space Kinetic Impact and Debris Branch (PL/WSSD)
Phillips Laboratory, Kirtland AFB, NM 87117-6008

ABSTRACT

A series of impact problems were analyzed using the Eulerian hydrocode CTH. The objective was to quantify the amount of energy dissipated locally by a projectile-infinite plate impact. A series of six impact problems were formulated such that the mass and speed of each projectile were varied in order to allow for increasing speed with constant kinetic energy. The properties and dimensions of the plate were the same for each projectile impact. The resulting response of the plate was analyzed for global KE, global momentum and local maximum shear stress. The percentage of energy dissipated by the various HVI phenomena appears as a relative change of shear stress at a point away from the impact in the plate.

INTRODUCTION

A hypervelocity impact (HVI) between two bodies of large relative size is characterized by severe localized phenomena such as material phase change, jetting and high strain rate. Until the impacting bodies reach this relative velocity range (7 to 12 km/s), these phenomena will not be as pronounced. In this case, the energy deposited by the projectile will not have these avenues of dissipation. In this paper, we have divided the process of HVI into an early-time and a late-time phase. This division will facilitate the understanding of the kinematics leading from the early-time phase to the late-time phase where breakup of a satellite is most likely.

The early-time phase includes those effects which occur during the first 3 to 4 μs of a hypervelocity impact (projectile speed greater than 5 km/s). The kinematics of the early-time phase are restricted to the area of impact. Global response does not yet occur at this time especially for large relative size difference between the target and the impactor.

The late-time phase occurs when the stress waves induced by the impact have had a chance to propagate outward to a distance of about 16 diameters of the projectile. At this time, the phenomena occurring at the impact site have begun to stabilize. Also, at this distance and further from the area of impact, the deformations are not as severe and there is certainly no hydrodynamic behavior of the material.

The quantity and rate of energy deposition required to cause the breakup of an orbiting spacecraft is a function of the impactor's speed, mass and shape. An understanding of the deposition and dissipation of energy in a structure from an impact will allow us to understand and predict the post-impact propagation of energy-containing stress waves throughout the structure at latter times. Therefore, the scope of this study does not seek to explain the early-time phenomena, but rather the

effect that these local phenomena have on the overall global response. The severity of the global response to a given amount of KE deposition will depend on the localized early-time dissipation of this KE. An example of a well-known localized dissipative effect absorbing much of the energy of an impact is the collapse of voids in a porous material [2]. Porous materials are used as shielding since voids in the material collapse and absorb much of the energy deposited by an impact. Therefore, less damage-producing energy is left to propagate throughout the structure.

PROBLEM FORMULATION

The 2 and 3-D hydrodynamic code CTH [1] was used to model a series of projectile impacts onto a plate. The speed and mass of the projectiles were varied but KE was held constant. The series of two-dimensional models which were set up had the initial conditions shown in Table 1. The projectiles for the six cases were represented with a circle which can be thought of as an infinite cylinder in three dimensions. The projectiles impacted in a direction normal to the plate surface. Since the models are two-dimensional, mass, kinetic energy and momentum are on a per unit length basis. The location on the mesh of the plate and projectile for case 1 is shown in Figure 1. The left boundary of the mesh ends at 5 cm although the plotting program extended
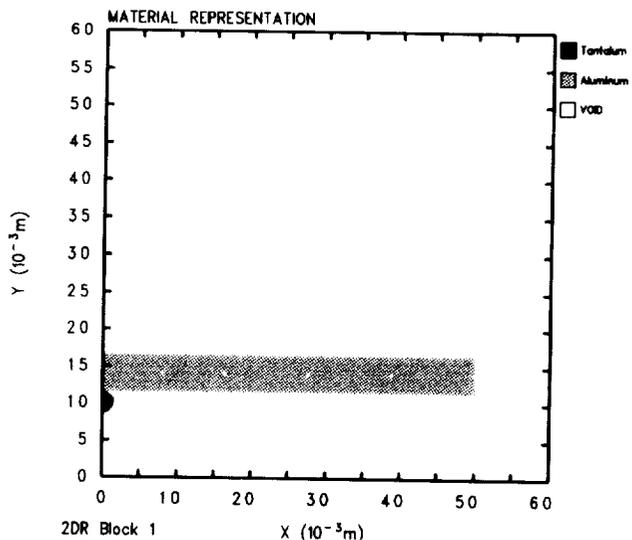
Figure 1. Model configuration of Case 1 at t=0.0

C-5

Projectile	Diameter (cm)	Speed (km/s)	Mass (g)	Kinetic Energy (kJ)	Momentum Kg-m/s (X-axis)
1	0.15	14.0	1.178	115.45	16.49
2	0.35	6.00	6.414	115.45	38.48
3	0.55	3.82	15.84	115.45	60.48
4	0.75	2.80	29.45	115.45	82.47
5	0.95	2.21	47.25	115.45	104.46
6	1.15	1.83	69.25	115.45	126.45

Table 1. Initial Conditions.

	Projectile	Target
Material:	Elemental tantalum	Elemental Aluminum
Shape:	Cylinder	Plate, 0.5 cm thick
Density:	16.6667 g/cm^3	2.713 g/cm^3
Equation of state:	Mie-Gruneisen	Mie-Gruneisen
Fracture model:	Max pressure criteria	Max pressure criteria
Strength model:	Elastic-plastic	Elastic-plastic

Table 2. Physical properties of the target and projectile models.

the plot frame to 6 cm. The points along the mid-section of the plate are Lagrangian tracer points at which stress was monitored. A plane of symmetry or 'mirror image' was defined along the left side of the mesh while the right side of the mesh was defined as a nonreflective boundary. These boundary definitions allow the modeling of an infinite plate. The mesh cell size was 0.5 mm x 0.5 mm (2-D cartesian). The physical properties of the plate and projectile are listed in Table 2. Tantalum was chosen as the material for the projectile because of its high yield and fracture strengths. For the scope of this study, it was best that the projectile remain as intact as possible since we were only interested in the response of the plate.

NUMERICAL RESULTS

The projectile in Case 1 had the highest speed and was the only one to induce localized phase changes. Figure 2 shows the progress of the impact for Case 1 at 1.27 μs. The darkly shaded region represents the material of the projectile and the plate material is shown in a lighter shade. It appears in Figure 2 that part of the aluminum plate has migrated onto the path of the projectile after it has passed through. This is explained by referring to Figure 3 which is a density representation of the same projectile-plate model at the same time (1.27 μs) as in Figure 2. The spacing of the dots in Figure 3 is proportional to the density of the material as shown by the legend on the upper right corner of the plot. From Figure 3 it can be seen that the material behind the projectile, directly in its path, is of much

lower density than the solid aluminum making up the rest of the plate. This indicates that a phase change of the aluminum is occurring here due to the tremendous amount of kinetic energy deposited at this local.

The occurrence of phase change is further substantiated by Figure 4. The projectile-plate model is overlaid with maximum shear stress contours. Each contour represents a region of constant shear stresses which are quantified in the legend on the right. The darkest region represents the materials of the plate and projectile. The lightest contour represents the lowest value of stress among the eight contour values. It is seen from Figure 4 that the stress concentration on the aluminum plate is decreasing in the direction towards the impact site. After the lightest contour, there is no representable stress further on. This substantiates the fact that the material in this region has undergone a phase change and is no longer able to support shear stress waves since it is no longer solid. It will be shown below that a slower projectile impact at this same time will support shear stress waves closer in to the projectile-plate contact interface. Finally, phase change is also apparent by studying the particle velocity vectors at 1.27 μs in Figure 5. The random directions of the vectors behind the projectile along its path suggest the behavior of a non-solid material phase.

The decreased speed of the projectile in Case 2 did not induce localized phase changes. In this case, the aluminum material of the plate near the impact site did not migrate onto the path of the projectile (see Figure 6). The density representation at 1.26 μs is shown in Figure 7. The time of this plot (1.26 μs) corresponds closely to the time of the density plot of the preceding case (Figure 3). To be noted in Figure 7 is the uniformity of the density of the materials in the region of the impact. This indicates a lack of phase change occurring in the materials at this time. Also, unlike the previous case, shear stress waves can now propagate closer to the projectile-plate contact interface since the materials are still in a solid state (compare Figure 4 with Figure 8).

Kinetic energy versus time was calculated and plotted for the plate and projectile in each of the six cases (see Figures 9 and 10). The dropping off of each curve in these figures after a level section of constant KE is not a physical phenomena. This is caused by pieces of mass leaving the active mesh zone and therefore no longer accounted for in the energy calculations. A comparison of the curves of the KE of the projectiles with their respective plates indicates that energy is indeed being conserved. Figures 9 and 10 indicate that the faster and lighter projectiles are more effective at transferring their KE into the plate. However, we cannot conclude from these figures that the quantity of the late-time propagated energy will be greater for the smaller lighter projectiles. The KE calculation on the plate as a function of time is insensitive to concentrations of KE in the plate since the calculation is a summation over the entire plate material including unattached fragments within the mesh. Kinetic energy producing phase changes may be concentrated at the impact site and dissipated locally and may not propagate extensively. Since the initial KE of the projectiles is constant, the

difference in the amount of KE in the plate after impact should be due only to the behavior of the material in the projectile. Also, the projectile will carry away some KE after perforation. Perforation occurred for all six projectile impacts but at different times as indicated by the stabilization times of the KE curves in Figure 9. The amount of KE in the plate was conserved at a constant level after each projectile had perforated the plate.

The lighter and faster projectiles were also more effective at transferring momentum to the plate as shown in Figure 11. Figure 11 is a plot of the momentum of the center-of-mass of the projectile in the direction of its path versus time. Since KE is proportional to the square of the velocity and momentum is proportional to velocity, it was not possible to maintain a constant initial momentum for the projectiles while keeping KE constant (see Table 1). However, KE energy rather than momentum will affect the local region of impact to a greater extent than momentum. Wu and Simons make a statement in [2] which describe this phenomena, ". . . a projectile or particle beam will deliver both momentum and energy to a target. The target will absorb the momentum through an increase in its mass mean velocity, dV_i, but dV_i is generally so small that the associated kinetic energy $(dV_i)^2$ is much less than the energy of the impact. The target must dissipate the impact energy at the site of the impact load. This dissipation mechanism is generally the strain energy of the target volume." For this reason, KE was held constant since it plays the greatest role in affecting the local region of impact.

Phenomena occurring at the site of impact which dissipate energy such as strain, melting, and fragmentation will be present to a greater or lesser extent depending on the speed and mass of the projectile. For this study, we were suspecting that phase changes in the material would dissipate some of the initial KE. To distinguish the amount of dissipation occurring, a point midway between both surfaces and 5 cm from the center of impact was monitored for maximum shear stress. This was done for each of the six cases and plotted in Figure 12. The phase changes induced by the lightest and fastest projectile succeeded in dissipating much of this KE. It can be seen from Figure 12 that projectile 1 produced a shear stress at a point 5 cm from the center of impact that is considerably below the maximum amplitude of the shear stress induced by the impact of the other 5 projectiles. The maximum amplitude of the shear stress for the other 5 projectiles were approximately the same. This is expected since their initial KE was constant and phase changes were not observed.

CONCLUSION

The series of calculations outlined in this paper have demonstrated the occurrence of energy dissipation in the early-time phase and the subsequent decrease in shear stress at latter-times. We have seen that phase changes occurring at the site of impact dissipate KE to such an extent that maximum shear stress is decreased by about 16 percent over slower projectiles with equal amount of KE. Since future breakup modeling will necessarily incorporate the effects of the early-time phase in order to predict the effects of the late-time phase, it is necessary to understand this dissipation phenomena.
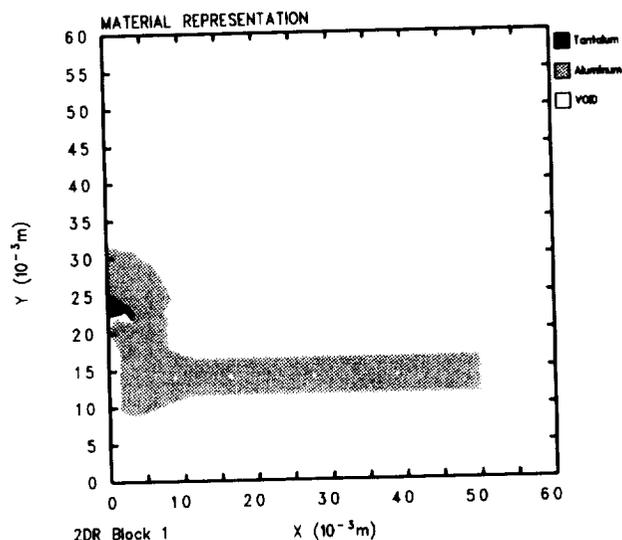
Figure 2. Model configuration of Case 1 at t = 1.27 μs.
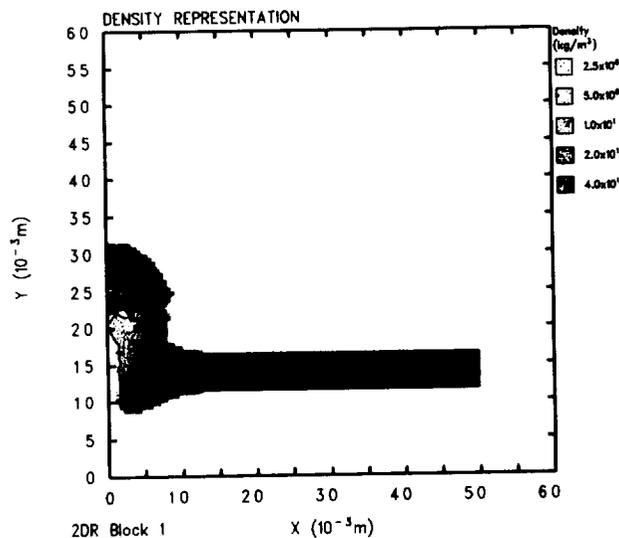
Figure 3. Density representation of Case 1 at t = 1.27 μs.

752
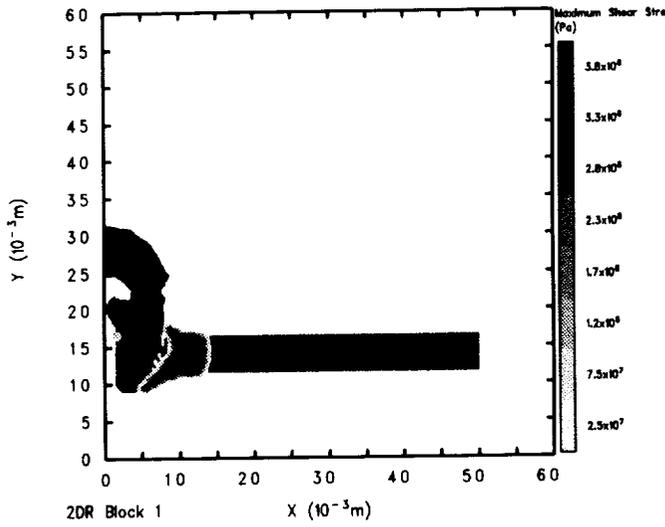
Figure 4. Maximum shear stress contour representation
of Case 1 at t = 1.27 μs.
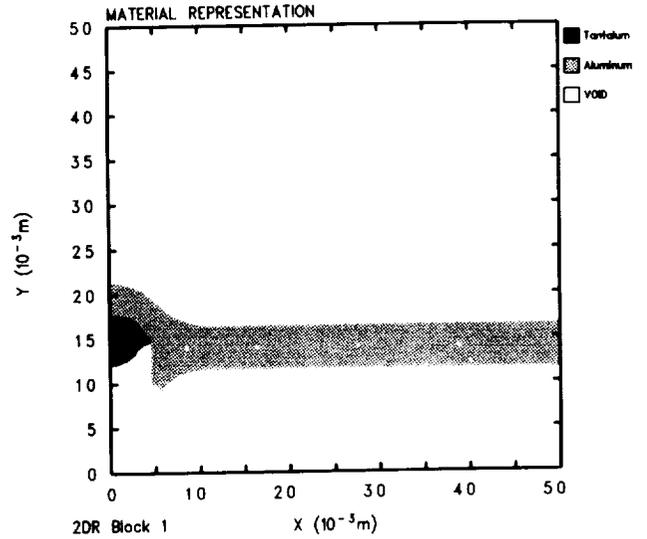
Figure 6. Model configuration of Case 2 at t = 1.26 μs.
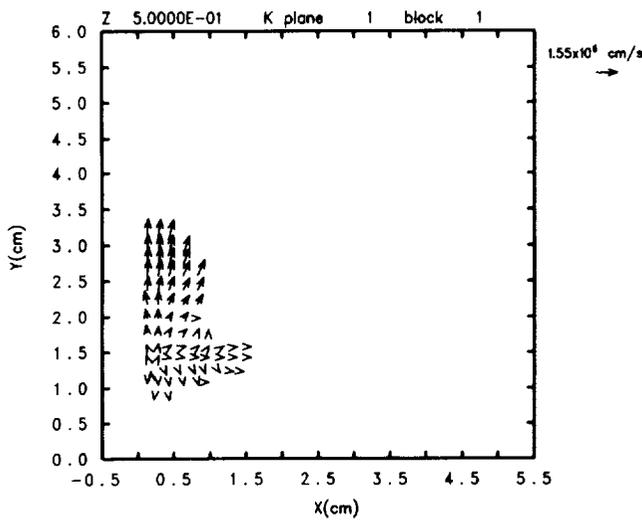
Figure 5. Velocity vector representaion of Case 1
at t = 1.27 μs.
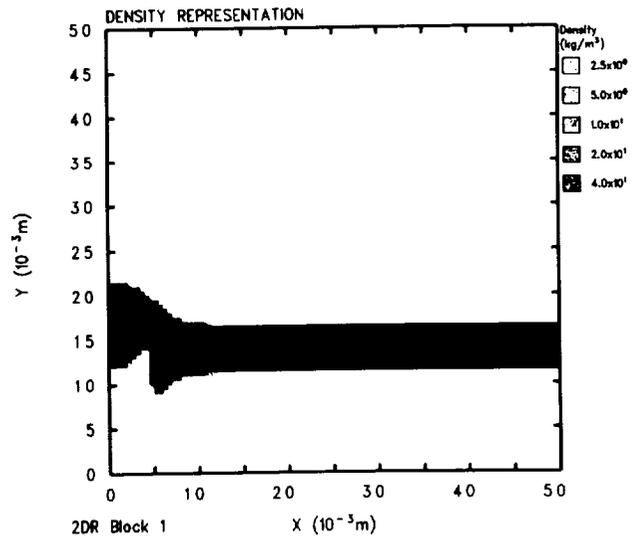
Figure 7. Density representation of Case 2 at t = 1.26 μs.
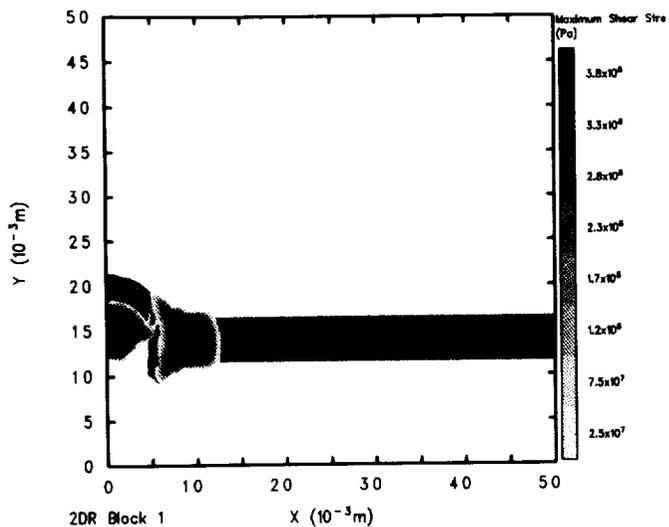
Figure 8. Maximum shear stress contour representation
of Case 2 at t = 1.26 μs.
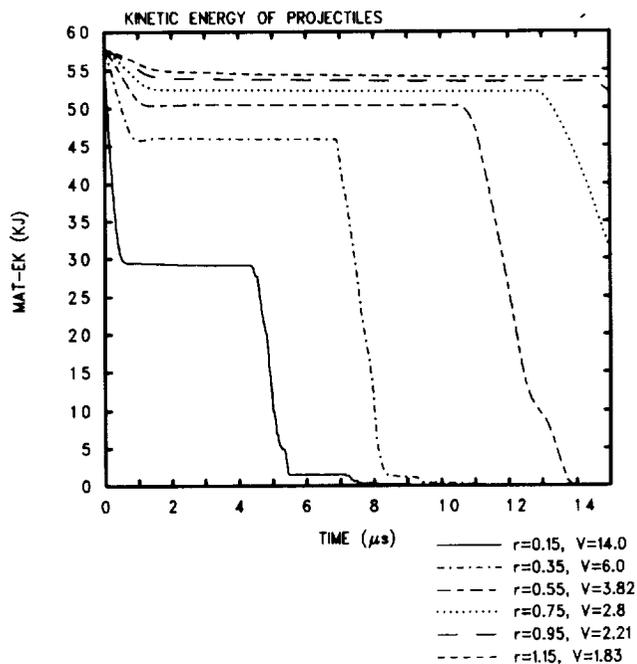
Figure 10. Total kinetic energy of the projectiles
for the six cases.
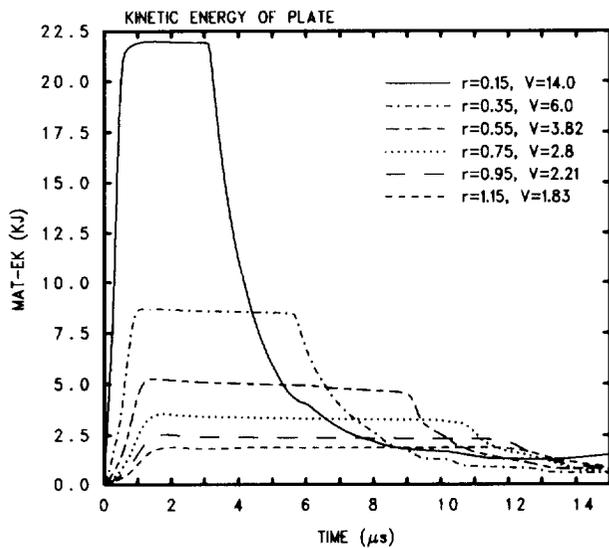
Figure 9. Total kinetic energy of the plate for the
six cases.
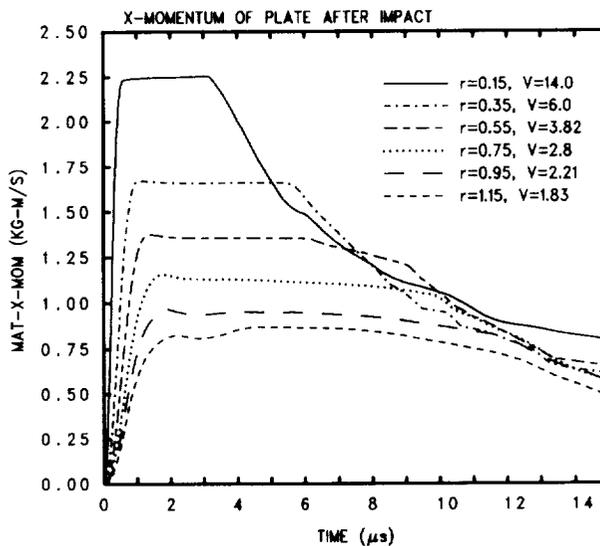
Figure 11. Momentum of plate after impact for the
six cases.
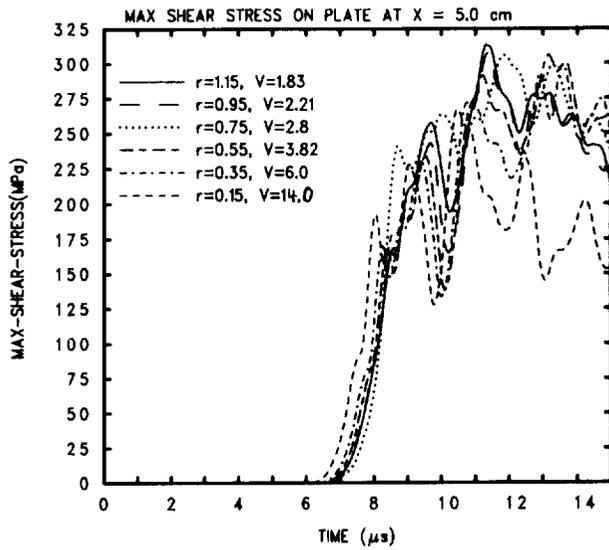
Figure 12. Maximum shear stress at 5 cm from center
of impact for the six cases.

REFERENCES

1. McGlaun, J. M., et. al., **CTH User's Manual and Input Instructions**, Sandia National Laboratories, Version 1.020, July 90.

2. Wu, K. S. and G. A., Simon, "Optimization of Target Mass Distribution Against Impulsive Loading," Physical Sciences Inc., Report PSI-2039/TR-750, January 88.

3. Thompson, S. L., and Lauson, H. S., **Improvements in the Chart D Radiation-Hydrodynamic Code III: Revised Analytic Equations of State**, Sandia National Laboratories, Report SC-RR-71 0714, June 1985, pp. 62-65.

Title: SPACE DEBRIS MEASUREMENT PROGRAM AT PHILLIPS LABORATORY

Authors: Phan D. Dao and Lt Ross T. McNutt
Phillips Laboratory, Geophysics Directorate.

Abstract:

Ground-based optical sensing has been identified as a technique for measuring space debris complementary to radar in the critical debris size range of 1 to 10 cm. The Phillips Laboratory is building a staring optical sensor for space debris measurement and considering search and track optical measurement at additional sites. The staring sensor is implemented in collaboration with Wright Laboratory using the 2.5 m telescope at Wright Patterson AFB, Dayton, Ohio. The search and track sensor is designed to detect and track orbital debris in tasked orbits. Progress report and discussion of sensor performance and search and track strategies will be given.

Session E6: SPACE DEBRIS - II: LABORATORY SIMULATIONS

Session Chair: J. C. Kolecki

DEBRIS AND MICROMETEORITE IMPACT
MEASUREMENTS IN THE LABORATORY

J. Resnick*, J. Grun, J. Crawford, R. Burris*, C.K. Manka, J.L. Ford, and B.H. Ripin
Space Plasma Branch, Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

*Physical Sciences Inc., Alexandria, VA

INTRODUCTION

The damage potential to spacecraft by micro-meteorites and orbital space debris has been recognized for some time. Damage caused by micrometeorites has been seen on the Solar Max Mission[1] and the LDEF Satellite[2] as well as other missions. Theoretical models (Kessler et al.[3]) suggest that most of the space debris in lower earth orbit (LEO) consists of particles smaller than 100 μm that travel at speeds of 10-15 km/sec (See Figure 1). Micrometeorites tend to move faster, at 20 km/sec, but are fewer in number compared to space debris.

Since opportunities to retrieve and study space-exposed material are rare, we have developed a method to simulate space debris in the laboratory. This method, which is an outgrowth of research in inertial confinement fusion (ICF), uses laser ablation to accelerate material. Using this method we have accelerated single 60 μm aluminum spheres to 15 km/sec and larger 500 μm aluminum sphere to 2 km/sec. Also, many small (<10 μm diameter) irregularly shaped particles have been accelerated to speeds of 100 km/sec. Studies similar to ours have been reported by Borodziuk and Kostecki.[4]

LASER ABLATION

In the laser ablation method of accelerating particles, a powerful 5 ns duration laser
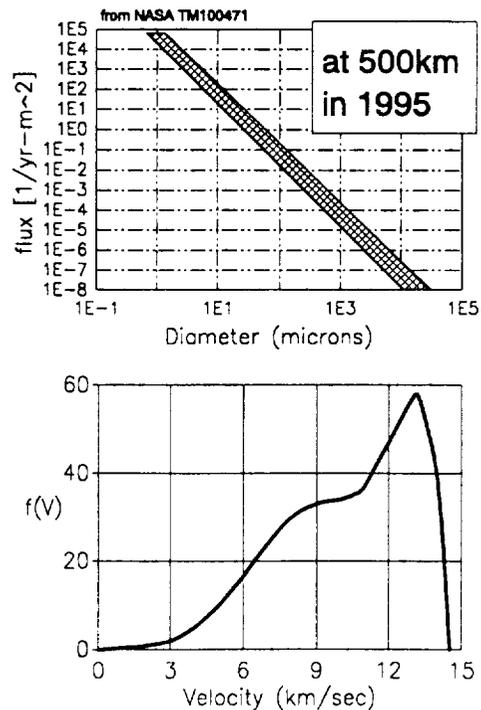
Figure 1: Size and Velocity of Micrometeorites in lower earth orbit

pulse is focussed with an irradiance of 10^{13} W/cm^2 onto the surface of a projectile. This irradiation heats the surface of the particle to temperatures of a few hundred eV, causing the surface to evaporate very rapidly. This rapid evaporation of the surface is called laser ablation. Generally, a few microns of surface material are ablated and the speed of the ablated material is 100 - 800 km/sec. This ablated material, which behaves like the exhaust of a rocket, accelerates the projectile (see Figure 2). Using energy and momentum conservation equations, the velocity of the projectile and the pressures on the surface can be calculated[5]. For small amounts of ablated mass (ΔM), the velocity of the projectile (v), the ablation velocity (u), and the mass of the projectile M_0 are related by

$$v/u = \Delta M/M_0 \qquad (1)$$

Thus, for an ablation velocity of 700 km/sec where only 2% of the mass is ablated, a particle velocity of 15 km/sec is achieved. The pressure on the surface of the accelerated particle using this technique is enormous.

Extensive theoretical and experimental studies of this process have been done for many years in the ICF community.[6,7] A theoretical study of the use of laser ablation for accelerated small (100 μm and smaller) spheres has been performed by Goela and Green.[8]

THE EXPERIMENT

The experimental set-up is shown in Figure 3. The Nd-glass (1.054 μm wavelength) Pharos III laser was focussed onto the surface of the projectile mounted on a holder (see Figure 4). Laser energies of 50 to 300 J were used with a pulsewidth of 5 ns. The projectile, accelerated by the ablation from its read surface, collided with a flat target (usually glass) up to 3 cm away from the holder. The projectiles were 60 to 500 μm diameter aluminum spheres. Shots could be repeated up to once per hour.

$$\text{Velocity} = \text{Vel. (ablated)} \cdot \frac{\text{Mass Ablated}}{\text{Mass Original}}$$
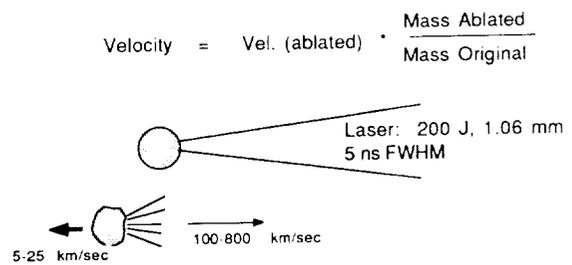
FIGURE 2: ACCELERATION BY LASER ABLATION
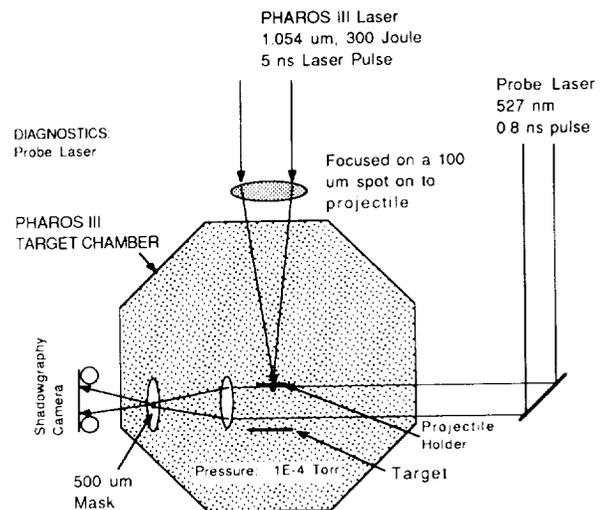
NOT TO SCALE

FIGURE 3
MICRO-METEORITE EXPERIMENT

To measure projectile velocities, dual-time dark-field shadowgraphy was employed.[9] This gave an image of the density gradients in the region surrounding the projectile and an image of the absorption region. In our experiment a sub-nanosecond pulsed 527 nm probe laser was directed across the path of the projectile. With two snap-shots (180 ns apart) of the projectile and debris as they moved away from the projectile holder, a velocity was calculated (see Fig. 5). Unfortunately, the electron density near the projectile due to the surrounding plasma is too high (above 10^{20} cm^{-3}) to see the outline of the projectile clearly. However, the plasma surrounding the projectile had little

759
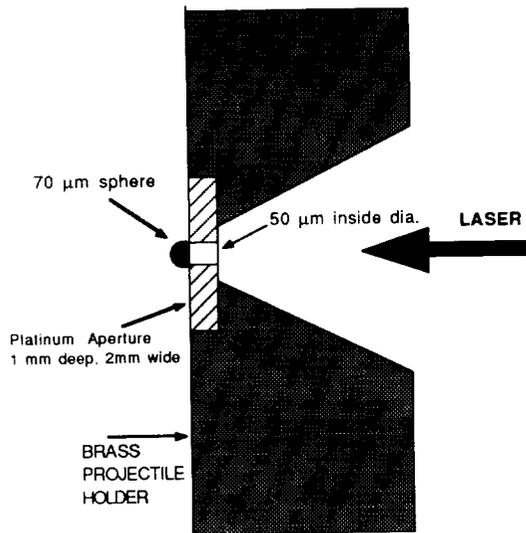
Figure 4

Projectile holder

70 μm sphere

50 μm inside dia.   **LASER**

Platinum Aperture
1 mm deep, 2mm wide

BRASS
PROJECTILE
HOLDER
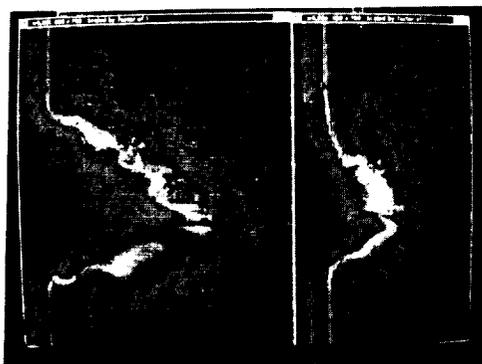
t + 180 ns   t

FIGURE 5

**Velocity measurement through
dark-field shadowgraphy**

mass and therefore, little if any effect on the target being impacted (this can be seen from the crater size and the lack of contamination). We assumed that the speed of the projectile within the plasma and the leading edge of the plasma seen is the same. This is supported by the fact that our calculation of projectile velocity and the measured velocities agree. We believe the margin of error for this measurement is approximately +/- 10%. A better velocity measurement will be made in the future using x-ray probe that can pass through the plasma surrounding the projectile but is stopped by the projectile.[10]

RESULTS

Half-millimeter diameter aluminum spheres were accelerated to 2 km/sec and produced single craters on a glass slide having a layer of copper foil on the front. The copper foil was needed to prevent the glass from shattering and make it easier to view any contamination by the aluminum sphere or the surrounding plasma. The craters produced had diameters of 1.5 mm and little contamination by aluminum outside the crater was seen.

The crater diameter was 1.5 times the crater depth. Figure 7 shows a crater made by a 250 μm aluminum sphere on a similar target. In one instance a 140 μm aluminum sphere was shot into a 1/8" (3 mm) thick piece of glass, breaking it in half. Aluminum spheres 70 μm in diameter were accelerated to speeds of 15 km/sec. Figure 8 shows the resulting 1 mm diameter crater.

Table I shows the size, mass, and speed achieved for the different sized aluminum spheres used. To achieve the maximum speeds listed in Table I laser energies of 300 J were needed. Laser energies greater than 300 J did not increase the projectile velocities. At projectile speeds greater than 8 km/sec some melted aluminum was observed within the crater. At lower speeds very little material was seen within the crater Plastic projectiles of similar sizes were also used and found to have similar speeds as the same sized aluminum projectiles.

760

Figure 6: Crater in copper thin on glass from a 500 μm Al sphere at 1 km/sec.

<----------------------->
5 mm

Figure 8: 1 mm diameter crater on glass from a 70 μm Al sphere at 15 km/sec.

<----------------------->
1.25 mm

Figure 7: Crater on glass from a 250 μm Al sphere at 7 km/sec.

<----------------------->
5 mm
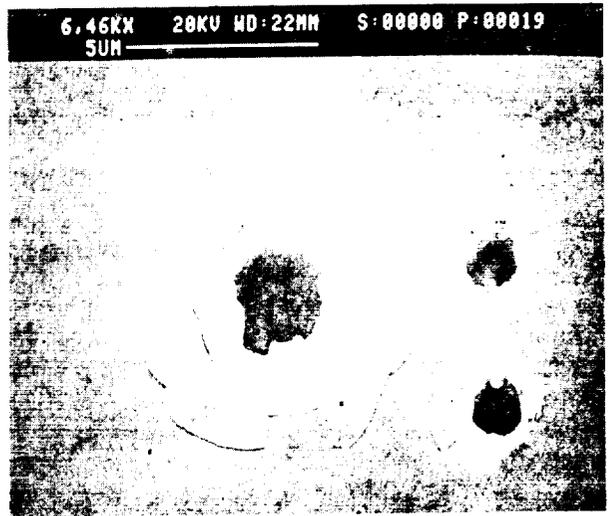
Figure 9: Examples of craters from irregular particles smaller than 10 μm traveling at speeds greater than 20 km/sec

761

TABLE I PROJECTILE PARAMETERS--
Aluminum

Projectile Diameter (μm)	Mass (μg)	Velocity Achieved
70	0.5	15 km/sec
140	4	12 km/sec
250	22	8 km/sec
500	177	2 km/sec

Thin foils 10 μm thick and 1 mm wide were accelerated using laser ablation. These foils broke up very quickly into many irregularly shaped particles smaller than 10 μm. These fragments impacted the targets at speeds of 20 to 100 km/sec. The greater speeds were possible due to the small areal mass of the material being accelerated. Figure 9 shows some of the craters formed on aluminum by such particles.

CONCLUSIONS

Acceleration of 70 μm aluminum and plastic spheres to speeds of 12 to 15 km/sec has been accomplished. Smaller irregularly shaped particles have been accelerated to speeds well above 20 km/sec. Given the ability to vary such parameters as projectile velocity, size, and incident angle, as well as the target and projectile materials, this technique should prove very useful for future studies of micrometeorite damage as well as hypervelocity impact (HVI) studies.

ACKNOWLEDGEMENTS

REFERENCES

1. J.L. Warren, et al., "The Detection and Observation of Meteoroid and Space Debris Impact Features on the Solar Max Satellite", Proceedings of the 19th Lunar and Planetary Science Conference, pp. 641 - 657, 1989.

2. D.S. McKnight et al., "Space Debris and Micrometeorite Events Experienced by WL 9829-9833" June, 1991.

3. D. J. Kessler et. al., "Orbital Debris Environment for Spacecraft Designed to Operate in Low Earth Orbit", NASA Doc. TM-100471, Sept., 1988.

4. S.R. Borodziuk and J.L. Kostecki, "Studies of hypervelocity impact problem by means of laser-target experiments-A new approach", Laser and Particle Beams (1990), Vol. 8, No. 1-2, pp. 241-245.

5. S.P. Obenschain, R.H. Lehmberg, and B.H. Ripin, Appl. Phys. Lett., Vol. 37, p 903.

6. B.H. Ripin et al., "Laser-plasma interaction and ablative acceleration of thin foils at $10^{12} - 10^{15}$ W/cm^2", Phys. Fluids, Vol. 23, No. 5, p. 1012, May, 1980.

7. J. Grun et. al., "Ablative acceleration of plasma targets to high velocities", Phys. Fluids, Vol. 26, No. 2, p 588, Feb., 1983.

8. J.S. Goela and B.D. Green, "Ablative acceleration of small particles to high velocities by focussed laser radiation", J. Opt. Soc. Am. B. Vol. 3, No. 1, p. 8 Jan. 1986.

9. J.A. Stamper et al., J. Appl. Phys., Vol. 52, p. 6562, 1981.

10. R.R. Whitlock et al., "Spectrally resolved flash x radiography of laser accelerated targets", SPIE Vol. 348 15th Int. Congress on High Speed Photography, 1983.

Simulating Hypervelocity Impact Effects on Structures
Using the Smoothed Particle Hydrodynamics Code MAGI

Larry Libersky
Center for Explosives Technology Research
New Mexico Institute of Technology
Socorro, N.M. 87801

Firooz Allahdadi
Phillips Laboratory
Space Kinetic Impact & Debris Branch
Kirtland AFB N.M. 87117

Theodore C. Carney
Advanced Sciences Inc.
6739 Academy Road N.E.
Albuquerque, N.M. 87109

INTRODUCTION

Analysis of interaction occurring between space debris and orbiting structures is of great interest to the planning and survivability of space assets. Computer simulation of the impact events using hydrodynamic codes can provide some understanding of the processes but the problems involved with this fundamental approach are formidable. First, any realistic simulation is necessarily three–dimensional, e.g., the impact and breakup of a satellite. Second, the thickness of important components such as satellite skins or bumper shields are small with respect to the dimension of the structure as a whole, presenting severe zoning problems for codes. Thirdly, the debris cloud produced by the primary impact will yield many secondary impacts which will contribute to the damage and possible breakup of the structure. Characterization of the debris cloud requires accurate fragmentation modeling as well as accurate tracking of the fragments through large regions of void. For these reasons hydrodynamic simulation of hypervelocity impact and breakup of orbiting structures is extremely difficult. We have approached the problem by choosing a relatively new computational technique that has virtues peculiar to space impacts. The method is called Smoothed Particle Hydrodynamics (SPH). In this paper we describe the SPH method and why we believe that it can be used to answer many questions concerning the survivability of space assets due to kinetic impacts. We also present several calculations to show the power of SPH towards such problems.

SPH BASICS

Smooth Particle Hydrodynamics is unique in computational fluid dynamics in that SPH uses no grid. It was the genius of the inventors, Lucy (1977) and Gingold & Monaghan (1977) to figure out how to take a derivative (get the force on a fluid element) without using a mesh. Previously a mesh was the only known way to compute a spatial derivative using finite differences. The mathematical theory of SPH will not be discussed here. The reader is referred to Gingold & Monaghan (1977,1982), Monaghan (1982,1985) and Monaghan & Gingold (1983) for detailed treatment of the subject. We present here only some basic features as discussed by Benz (1989) that are necessary to understand the method. Consider a function f, a kernel W which has a width measured by the parameter h, and the following equation:

$$< f(\mathbf{r}) > = \int W(\mathbf{r} - \mathbf{r}', h) f(\mathbf{r}') d\mathbf{r}'. \quad (1)$$

If we impose a normalization condition such that the integral of W is unity, then it follows that

$$< f(\mathbf{r}) > \overset{h \to 0}{\to} f(\mathbf{r}). \quad (2)$$

Relation (1) therefore defines the kernel estimate $<f>$ of f. If W is the Dirac delta function then we have the equality $<f>=f$. Now suppose f is known only at N discrete points that are spatially distributed according to the number density distribution:

$$n(\mathbf{r}) = \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j) \qquad (3)$$

If the number density at $\mathbf{r}_j$ is written as

$$< n(\mathbf{r}_j) > = \frac{\varrho(\mathbf{r}_j)}{m_j} , \qquad (4)$$

thus introducing the concept of particle mass (m), the following equation can be derived:

$$< f(\mathbf{r}) > = \sum_{j} f_j \ W(\mathbf{r} - \mathbf{r}_j, h) \frac{m_j}{\varrho_j} \qquad (5)$$

This equation defines a procedure for transforming integral equations to particle equations and is therefore called "integral evaluation by the particle method." The choice of kernel or "smoothing function" is discussed by Monaghan and Lattanzio (1985). The W most frequently used in SPH codes is a B-spline with compact support which goes to a zero at a distance $2h$ from its peak. When the conservation laws of fluid dynamics are cast into the SPH framework using the procedure described above, the following equations are obtained.

$$\varrho_i = \sum_{j} m_i W_{ij} \qquad (6)$$

$$\frac{d\vec{u}_i}{dt} = -\sum_{j} \left[\frac{P_i}{\varrho_i^2} + \frac{P_j}{\varrho_j^2} + \Pi_{ij} \right] m_j \ \vec{\nabla}W_{ij} \qquad (7)$$

$$\frac{de_i}{dt} = \frac{P_i}{\varrho_i^2} \sum_{j} m_j(\vec{u}_i - \vec{u}_j) \cdot \vec{\nabla}W_{ij}$$
$$+ \frac{1}{2} \sum_{j} m_j \Pi_{ij}(\vec{u}_i - \vec{u}_j) \cdot \vec{\nabla}W_{ij} \qquad (8)$$

Equation (6) is the density computation of particle "i" using the masses of neighboring "j" particles. The acceleration of particle "i" is given by (7) and the evolu-

tion of the specific internal energy (e) is described by equation (8). These equations also involve the pressure (P) and the artificial viscous pressure Π. Terms involving material strength are omitted here but are discussed by Libersky and Petschek (1991). It should be emphasized that equations (6) thru (8) are the continuum equations of fluid dynamics cast into a discrete Lagrangian frame by kernel interpolation. Therefore, the Smoothed Particle Hydrodynamics method is a true and complete hydrodynamic calculational procedure.

IMPORTANCE TO SPACE IMPACTS

Simulating the mechanics of irreversible processes that take place between a structure and a projectile during a hypervelocity collision in space is extremely difficult. There are three main reasons for this. First, problems are three-dimensional. Second, the thickness of important components such as satellite skins, electronic components or bumper shields are small with respect to the dimension of the structure as a whole. Three-dimensional calculations using mesh-based codes do not seem feasible for such problems. Thirdly, the debris cloud produced by the primary impact produces secondary and tertiary impacts important to the overall damage and breakup. Simulation of these events requires detailed characterization of the debris cloud which in turn requires good fragmentation modeling as well as accurate tracking of the fragments through large regions of void. Eulerian codes have difficulty tracking sub-grid scale fragments through the mesh. Also, large regions of void within the structure need to be zoned in anticipation of material arriving there at some later time. SPH suffers from neither of these difficulties because there is no mesh. Following the debris cloud through large regions of void presents no difficulty to SPH. Furthermore, interfaces between materials in a problem consisting of several materials are accurately tracked. These virtues of SPH are true of any Lagrangian code, except that mesh-based Lagrangian codes cannot treat large fluid distortions. Obviously, hypervelocity impact produces highly distorted flows. It can be said

764

that SPH contains the best features of grid–based Eulerian and Lagrangian methods without the limitations of either. The price to be paid is that SPH is slower than the other methods, (having to determine new neighbors each computational cycle) but not prohibitively so. The efficiency of SPH appears to be much better than Eulerian in 3D (Durisen, 1988) and we are paying close attention to vectorization and parallelization. The ability of SPH to accurately track debris resulting from hypervelocity impact through large regions of void in 3 dimensions and compute the impact of the debris on other parts of the structure make the method extremely attractive for space applications. In working towards that goal we have performed several two-dimensional calculations to evaluate SPH and get a feel for how impact events might damage and lead to breakup of the large-scale structure.

SYNTHESIS OF A PREDICTIVE BREAKUP MODEL

Determination of processes that contribute to the failure and ultimately to the breakup of a complex and integrated space asset under impulsive loading is of vital interest. Currently, such a predictive model does not exist. Until recently, virtually all breakup modeling of a spacecraft under intensive loading used a phenomenological approach. Although this approach provides some qualitative measure of the interactions, it lacks the physics necessary to identify those processes that control generation of the resulting debris cloud. To characterize the environment of a debris cloud accurately (in terms of debris mass, velocity and number distribution), a first principles physics based predictive model has been synthesized. This approach considers the synthesis of a total predictive model based on the response of elementary components. Internal components of a generic satellite are shown in Figure 1 and an example matrix of calculations leading to the total predictive model is shown in Figure 2. The design of these and other calculations is motivated by the various geometries of the components internal to the satellite. Two simulations of local impact events on small regions of a large scale structure are shown in Figure

3 and Figure 4. These calculations give us detailed understanding of how the debris cloud interacts with nearby structural elements producing damage.

Figure 3a is a particle plot showing initial conditions for the impact of an tantalum projectile on an aluminum (2024–T86) frame. Each arm of the frame had length 20.5 cm and thickness 0.5 cm. The projectile had 1.0 cm sides, a speed of 7 km/s and a 60 degree impact obliquity. The calculation was performed in a two–dimensional Cartesian frame of reference. A Gruneisen equation of state and an elastic-perfectly plastic constitutive model were used to describe the metals. The calculation used 12,240 particles and the smoothing length was 0.08333 cm. Results of the calculation at 1 and 20 microseconds are shown in Figures 3b and 3c respectively. A very large opening is created in the first plate impacted due to the large impact obliquity. This "hole" is approximately 7 times the initial projectile size. One end of the plate is bent inward. There is a large splash of material moving upward and away from the structure typical of high speed cratering events and a debris cloud expanding towards the other arm of the structure. This cloud has fragments of various sizes but most of the mass is concentrated in the part of the cloud that is about to impact the second structural arm.

Results of a similar calculation employing a slightly more complex aluminum structure are displayed in Figure 4 where particle plots at four different times (0, 10, 20, 30 μs) are shown. In this calculation the debris cloud, resulting from the impact upon one member of the aluminum frame, is seen impacting other structural elements. Severe damage is seen on the upper panel. In fact, a secondary debris cloud has been produced by the interaction. It is easy to see from this calculation, that in an actual satellite with many more composite parts, how a cascade of debris clouds could form to cause massive breakup of the structure. Some damage is also seen on the outside panel furthest from the impact. These calculations were performed on a 1 megaflop machine and required approximately 1 hour of cpu time per 10 μs of simulation time.

Target Name: chucksat.tp4
Aimpoint(X,Y,Z) (cm) = (0.00 0.00 45.50)
Observer Azimuth (degrees) = 91.00
Observer Elevation (degrees) = 0.00
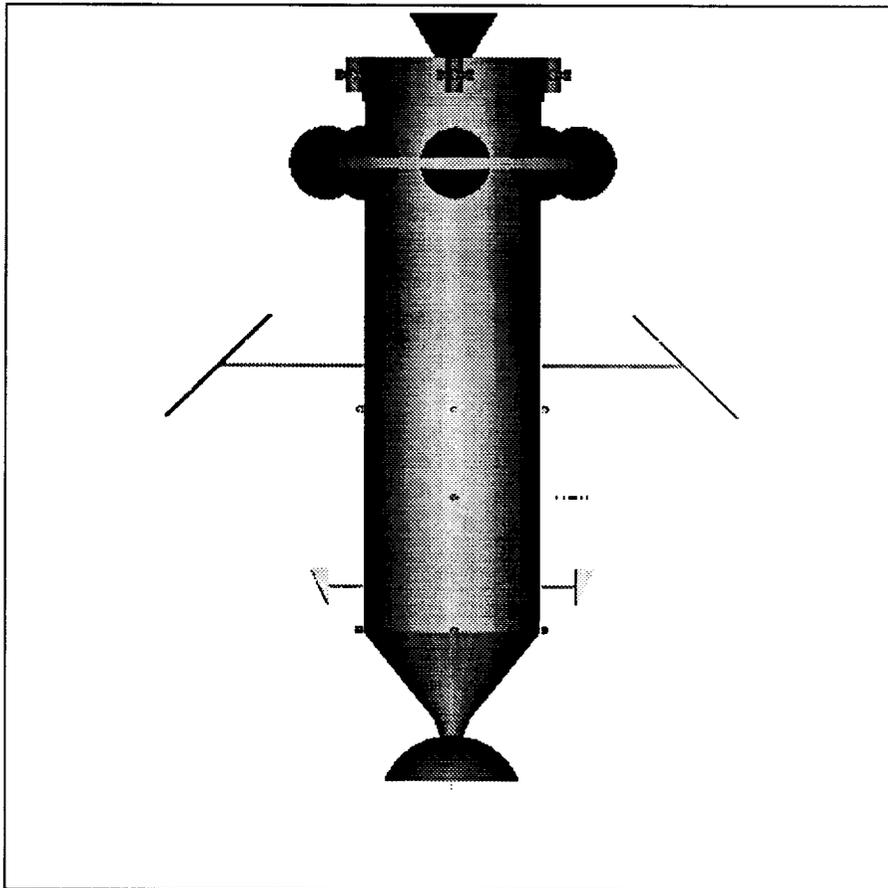
Figure 1. A Generic Satellite

766
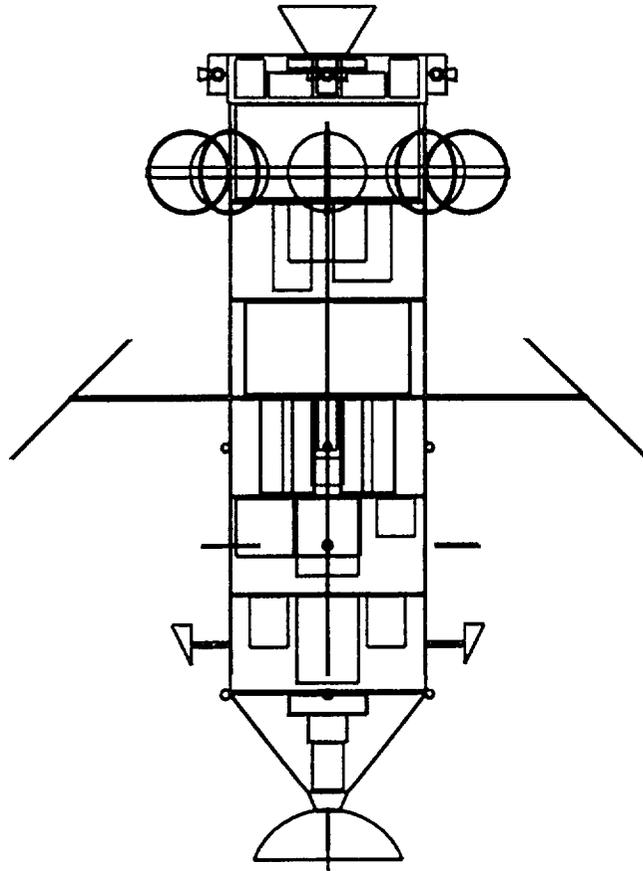
Figure 1A. Interior Components of Satellite

Matrix of Calculations Leading to a
## Satellite Breakup Model
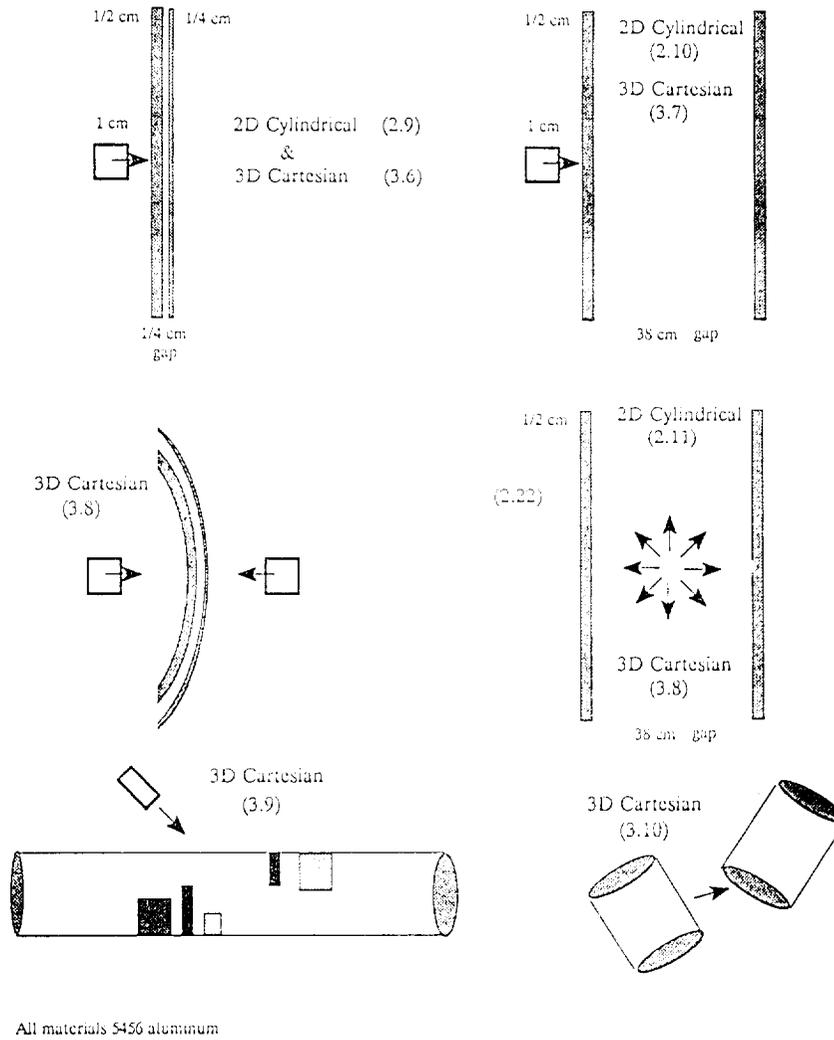
Figure 2. Matrix of Calculations

(3a)   Initial configuration.
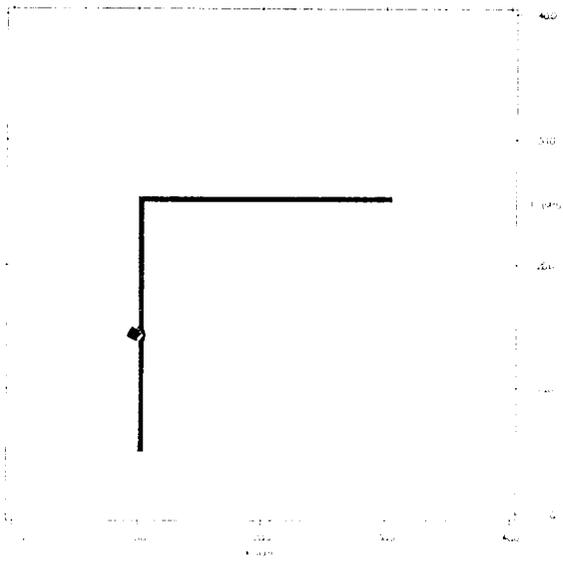

(3b)   1 microsecond after initial impact.

Figure 3.   Material plot of a tantalum projectile penetrating an aluminum
            frame at 60° obliquity.
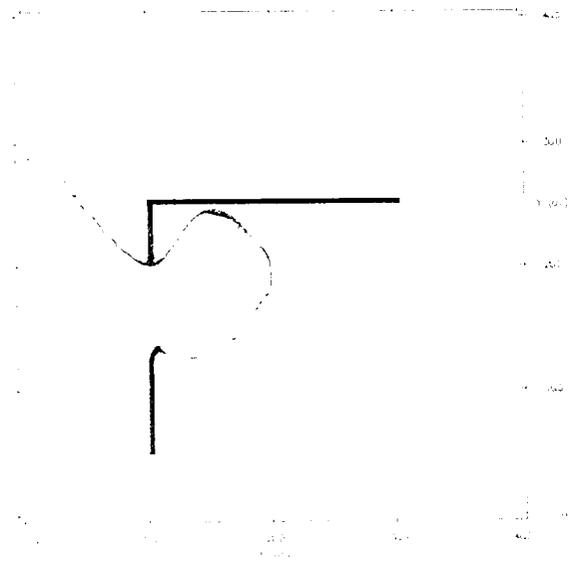
(3c) 20 microseconds after initial impact.

Figure 3 (cont).   Material plot of a tantalum projectile penetrating an
                    aluminum frame at 60° obliquity.
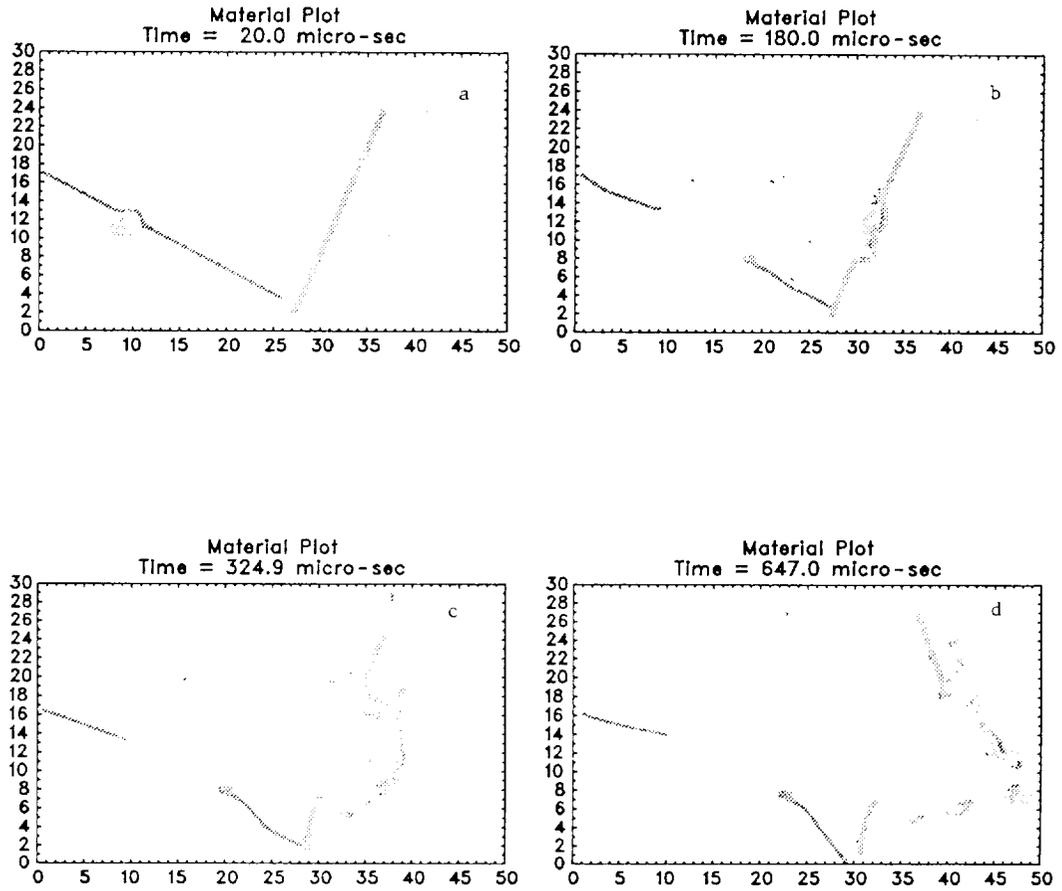
Figure 5. Smoothed Particle Hydrodynamics simulation of a liquid filled steel pellet impacting multiple spaced plates at 2 km/s and high obliquity. The first plate impacted is aluminum amd the remaining two are steel. The calculation was done in two-dimensional Cartesian coordinates using the SPH code MAGI. Frames a - d show results at 20, 180, 325, and 647 μs.

Density (gm/cc)
Time = 3.0 micro-sec

a

2.0000E-01   1.8000E+00
Density (gm/cc)
Time = 30.0 micro-sec

b

2.0000E-01   1.8000E+00
Density (gm/cc)
Time = 60.0 micro-sec

c

2.0000E-01   1.8000E+00
Density (gm/cc)
Time = 120.0 micro-sec

d

2.0000E-01   1.8000E+00
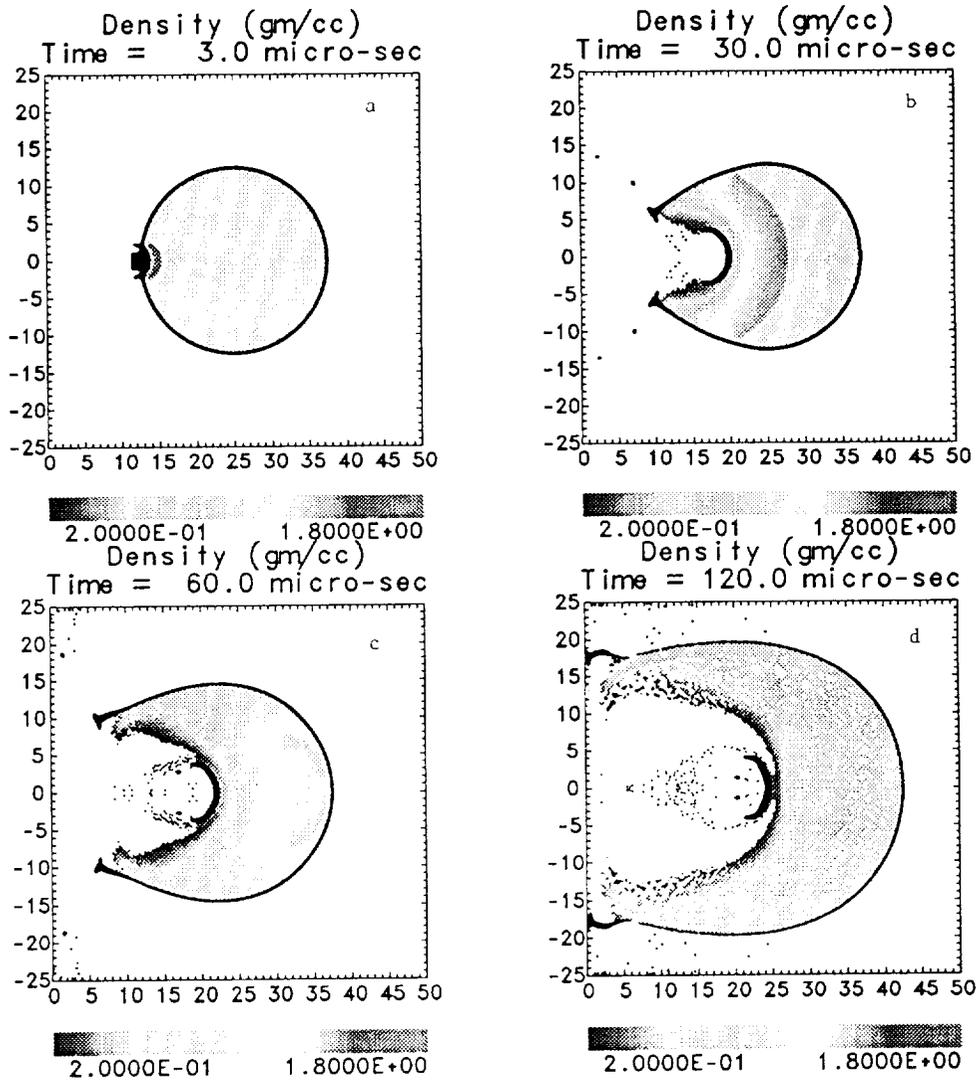
Figure 6. Smoothed Particle Hydrodynamics simulation of a liquid filled aluminum tank hit by an aluminum projectile moving at 7 km/s. The calculation was done in two-dimensional Cartesian coordinates using the SPH code MAGI. The liquid particles are gray-scaled by density and the aluminum particles are shown in black. Frames a - d show results at 3, 30, 60, and 120 μs.

Several calculations have been presented – two hyper-velocity impacts on simple structures, and two other impact problems possessing interesting features. We are encouraged by the results. The code has been extended to three-dimensions and we are currently performing test problems.

REFERENCES

Benz, W., Numerical Modeling of Stellar Pulsation: Problems and Prospects, Nato Workshop, Les Arcs, France, March, 1989 (Harvard Smithsonian Center for Astrophysics, Preprint No 2844, (1989).

Durisen, R.H., R.A. Gingold, J.E. Tohline and A.P. Boss, *The Astrophysical Journal*, 305, (1986), 281-306.

Gingold, R.A. and Monaghan, J.J., *Mon Not Roy Astron Soc* 181 (1977), 375-389.

Gingold, R.A. and Monaghan, J.J., *J Comput Phys* 46 (1982), 429-453.

Libersky, L.D. and A.G.Petschek, *Lecture Notes in Physics*, Springer Verlag, in print (1991).

Lucy, L.B. *Astron. J.* 83 (1977) 1013-1024.

Monaghan, J.J., *SIAM J Sci Stat Comput.*, 3 (1982), 422.

Monaghan, J.J., *Comp Phy Rev* 3 (1985), 71-124

Monaghan, J.J and Gingold, R.A, *J Comput Phys* 52 (1983), 374-389.

Monaghan, J.J. and Lattanzio, J.C., *Astron Astrophys* 149 (1985), 135-143.

ADDITIONAL IMPACT CALCULATIONS

Two additional calculations are presented in order to show the power of SPH towards impacts. Figures 5a–5d are particle plots showing a liquid filled projectile impacting a series of spaced plates at high obliquity. The projectile casing was steel and the simulated liquid was modeled as carbon–tetrachloride. The length of the projectile was 3.75 cm and the thickness 2.50 cm. The first plate to be impacted was aluminum with thickness 0.254 cm, and the remaining two plates were steel with thicknesses 0.48 cm and 1.58 cm respectively. The projectile speed was 1.6 km/s. The problem was run in a two–dimensional rectangular Cartesian frame of reference (plane–strain) with approximately 4,000 particles. The smoothing length was 0.127 cm. An elasctic–perfectly plastic constitutive model and Gruneisen equation of state were used to describe the solids. No explicit failure model was included in the calculation. Figure 5a shows the aluminium plate being impacted (t = 20 μs). Figure 5b shows the calculational result at 180 μs. The projectile has struck the first steel plate and the lower aluminum plate which was set in motion by the impact has also hit the steel and started to buckle. At 325 μs, as shown in Figure 5c, the aluminum plate has continued to buckle and the portion of the broken steel plate at the bottom of the figure has rotated as a result of the impact torque. These are two interesting features of the strength model operating in the code. It is very encouraging to see these effects captured by the code. Other interesting features are the crack formation at the back of the thick steel plate and the enhanced upward momentum of the first steel plate due to the crater splash from the projectile on the second plate. We expect these kinds of secondary effects to be important in the actual satellite breakup from impact. An obvious key feature of simulation is the "plug" of steel plate seen in Figure 5d. This is not due to any explicit failure model in the code but results from intrinsic model fracture in SPH. In response to elastic waves in the plate, some particles find themselves outside of the smoothing length range of communication with its

neighbors. This separation manifests itself as a "crack" which propagates through the metal. We cannot, at this point, claim that this effect is actually a physical one. Nevertheless, it appears that the code is trying, on its own, to accommodate failure, and the results seem physical. This calculation required 2 hours of cpu time on the CRAY 2 at the Phillips Laboratory.

Results of a hypervelocity impact (7 km/s) simulation of an aluminum projectile into a water–filled aluminum tank are shown in Figure 6 where the particles are gray–scaled according to their density. The diameter of the tank was 25 cm initially and the wall thickness was 0.5 cm. The projectile length was 3.0 cm and the thickness was 2.0 cm. The geometry was 2D Cartesian. Approximately 20,000 particles were used in the simulation. The smoothing length was 0.2 cm. Notice the shock propagation in the water and along the tank inner surface which comes to a focus at the rear of the cylinder and then strongly reflects back. The reflected shock is evidenced by the flattening of the particle distribution just ahead of the projectile. Notice also the rapid deceleration of the aluminum projectile and its large deformation. This calculation took 7 hours on the CRAY 2. For these two specialized problems there is no experiment data to which the simulations can be compared, so we must be careful not to draw unjustified conclusions about the codes performance. However, we can get a feel for how the code responds to difficult impact problems with the goal of extending the calulations to three–dimensions with extensive comparison to experiments.

CONCLUSIONS

The Smoothed Particle Hydrodynamics (SPH) approach to computational fluid dynamics has been briefly described with emphasis on its natural ability to model hypervelocity impact on orbiting structures. Our goal is to exploit these virtues of SPH towards the development of a complete structural breakup model in order to answer important questions concerning the survivability of space assets due to kinetic impacts.
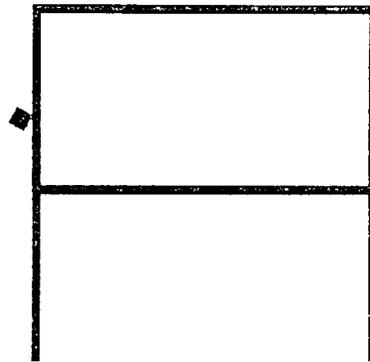
(4a)   Initial configuration.
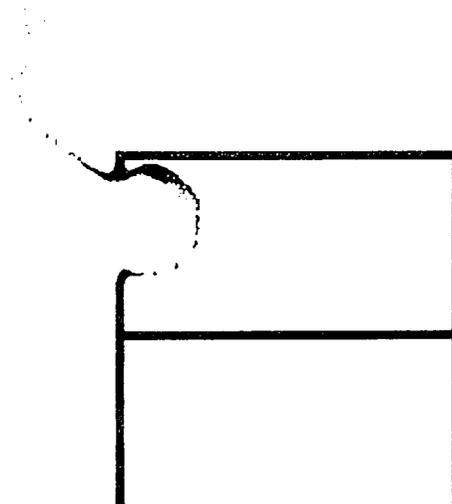
(4b)   10 microseconds after initial impact.

Figure 4.   Material plot of a tantalum projectile penetrating an aluminum
            A-frame structure at 60° obliquity.

775
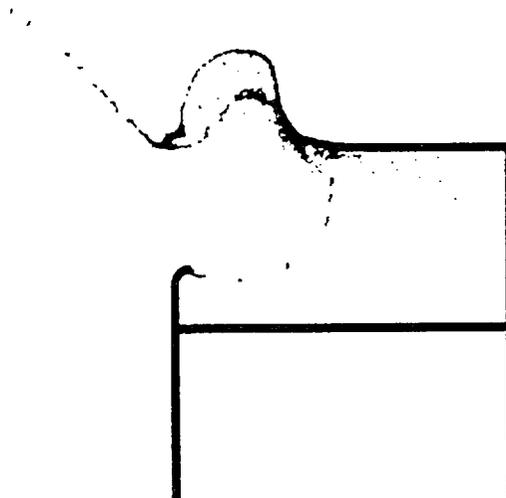
(4c)   20 microseconds after initial impact.
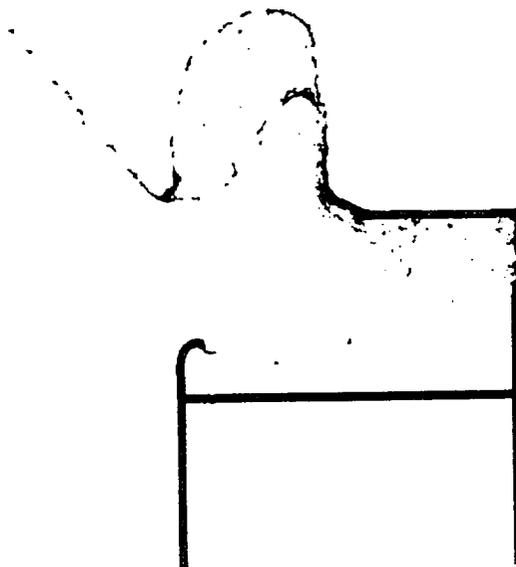
(4d)   30 microseconds after initial impact.

Figure 4 (cont).   Material plot of a tantalum projectile penetrating an
                    aluminum A-frame structure at 60° obliquity.

Session E7: SPACECRAFT INTERACTIONS

Session Chair: Lt. Col. Nelson

777

RADIATION-INDUCED INSULATOR DISCHARGE PULSES IN THE CRRES INTERNAL DISCHARGE MONITOR SATELLITE EXPERIMENT

A. R. Frederickson, E. G. Mullen,
D. H. Brautigam and K. J. Kerns,
Space Physics Division,
Air Force Phillips Laboratory, Hanscom AFB, MA 01731

E. G. Holman, Boston College, Chestnut Hill, MA.

ABSTRACT

The Internal Discharge Monitor, IDM, was designed to observe electrical pulses from common electrical insulators in space service. The sixteen insulator samples include G10 circuit boards, FR4 and PTFE fiberglass circuit boards, wires with common insulations, FEP teflon, and alumina. The samples are fully enclosed, mutually isolated, and space radiation penetrates 0.02 cm of aluminum before striking the samples. Pulsing began on the seventh orbit, the maximum pulse rate occurred on the seventeenth orbit when 13 pulses occurred, and the pulses slowly diminished to about one per 3 orbits six months later. After 8 months, the radiation belts abruptly increased and the pulse rates attained a new high. These pulse rates are in agreement with laboratory experience on shorter time scales. Several of the samples have never pulsed. If the pulses were not confined within IDM, the physical processes could spread to become a full spacecraft anomaly. The IDM results indicate the rate at which small insulator pulses occur. Small pulses are the seeds of larger satellite electrical anomalies. The pulse rates are compared with space radiation intensities, L shell location, and spectral distributions from the radiation spectrometers on CRRES.

INTRODUCTION

The purpose and background of the Internal Discharge Monitor (IDM) is documented in reference 1. Nearly twenty years ago it was determined that satellites were suffering from electrical discharge noise and EMP problems caused by the high energy electron space radiation belts. A number of conference sessions, symposia, and journal issues have been dedicated to the topic (Air force, NASA, IEEE, AIAA, ESA). Most of the previous work concentrated on the voltages of surfaces on the satellites. It became clear that the estimated surface voltages were not sufficient to cause the large number of events seen on satellites.

Space radiations acting directly on insulating materials should be able to produce electric pulses, independent of the potential of the spacecraft. The radiation intensity in space is so low that the pulse rates could not be predicted with any certainty. Ground testing, extrapolated to space radiation intensities, hinted that pulses might be as high as a few per day in space [1]. However, the extrapolation is very uncertain because the effects of electrical conductivity at high fields in the insulators can not be predicted in space service. The IDM is measuring the pulses produced by simple insulator structures which are isolated from the effect of the spacecraft potential itself. The IDM results are teaching us that simple insulators, irradiated by space electrons with energies predominantly above 100 keV, produce significant numbers of pulses. The results are also indicating that surface voltage may not be a prime cause of pulses.

The IDM instrument was designed and constructed by the Jet Propulsion Laboratory at the request of the U. S. Air Force. Briefly, the IDM exists to determine if spontaneous electrostatic discharges occur on insulators internal to spacecraft caused by

penetrating radiation. The IDM has been partially described in a previous publication [1] which reviews test results for IDM sample material under laboratory simulations of the Chemical Release and Radiation Effects Satellite (CRRES) space radiation environment. These test results provided valuable data for determining the final design of the instrument. This paper describes the IDM samples and some results from the first 100 days in space.

The samples were chosen for their generic nature; they are representative of the many materials and device structures used in most spacecraft. In addition, two samples are installed to measure in real time the possible response of the circuit boards of the MicroElectronics Package (MEP) which is adjacent to the IDM on the CRRES satellite.

As radiation (mostly electrons) penetrates the skin of a satellite and stops in the insulating materials, high voltages build up both on and in the insulator materials. Occasional spontaneous discharges of the insulators induce electrical pulses in sensitive satellite wiring. The magnitude of the pulses depends on many parameters including the area of charged insulator surface, the radiation spectrum, existing bias levels on adjacent circuits, the relative area of the sensitive circuit, the effective impedance of the sensitive circuit (as a function of frequency), voltages on other adjacent charged insulator surfaces, a complicated transfer function between the discharging circuit and the sensitive circuit, etc. The total process is much too complex for the IDM to study. The IDM only detects the number of pulses which occur above a threshold level on simple representative generic circuit elements. Thus, it tells us how often such pulses are occurring but nothing about the nature of the pulses. Based on present knowledge (ref. 1 and unpublished personal experience), the IDM is expected to experience pulse rates somewhere between zero and a few per day. The IDM results will significantly improve our knowledge of pulse rates in space.

The IDM is nearly a cubic foot in size, weighs 33 pounds, and uses 6.7 watts continuously. It was designed to last a minimum of two years in its environment. The CRRES orbit is roughly geosynchronous transfer with an apogee of 33,500 km, a perigee of 348 km, a period of 9 hours 52

minutes, and an equatorial inclination of 18.2 degrees. Thus the IDM is exposed to the inner belt for a short time, and to the outer electron belt for a long time. High energy electrons (> 150 keV) will be the dominant cause of IDM pulses.

SAMPLES

The samples were chosen after an extensive testing program as described in reference 1. Table I lists the samples and their configurations as flown on the IDM. The configurations are shown in figure 1. Most of the samples and configurations are also described in reference 1, however, configuration #1 is herein corrected (#1 in ref. 1 is not used in IDM) and configurations 7 and 8 are new.

SAMPLE CONFIGURATIONS
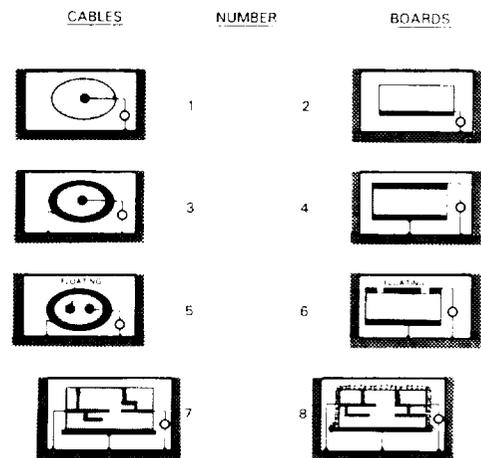
Figure 1. Sample Wiring Configurations. Heavy lines are electrodes, small circles represent the 50 ohm measuring circuit, and the curly line represents leaky paint. Each sample is faraday shielded from external fields by 0.2 mm aluminum on the top and 0.5 cm aluminum on the sides and bottom of its container. The feed-through [1] to the external detectors is composed of semi-insulating material which does not pulse. It is not shown in order to simplify the drawing.

TABLE I. DESCRIPTION OF IDM SAMPLES

V max is the maximum pulse voltage during ground tests, CONFIG is the number in figure 1 corresponding to the geometry of electrodes and sample, and PULSES is the number of pulses accumulated in the first one hundred days in space.

CHANNEL	SAMPLE DESCRIPTION	V max	CONFIG	PULSES
1	SC18 WIRE TYPE ET	1	1	9
2	TS TRIAX CABLE	5	5	0
3	MEP G10 SOLITHANE COATED ONLY	50	7	0
4	FR4 FIBERGLASS, 0.317 cm	5	2	41
5	RG 316 CABLE	0.5	3	0
6	ALJAC CABLE	1	3	0
7	ALUMINA, 0.102 cm	40	6	0
8	FR4 FIBERGLASS, 0.317 cm	1	4	0
9	FEP TEFLON, 0.229 cm	100	6	1
10	FEP TEFLON, 0.229 cm	0.2	4	0
11	PTFE FIBERGLASS, 0.229 cm	1	4	0
12	FR4 FIBERGLASS, 0.317 cm	5	2	12
13	FR4 FIBERGLASS, 0.317 cm	100	6	0
14	MEP G10 SOLITHANE WITH LEAKY PAINT	< 1	8	0
15	FR4 FIBERGLASS, 0.119 cm	0.25	2	39
16	PTFE FIBERGLASS, 0.229 cm	0.2	2	163

The maximum pulse voltage measured on a 50 ohm line (sometimes 25 ohms) is reported in Table I from previous tests: from reference 1 for configurations 1-6, and from reference 2 for configurations 7,8. Because the pulse voltage is so dependent upon many external factors, it is not the exact voltage which is of interest, it is the occurrence of a pulse of sufficient energy to interrupt normal circuit signals that is of interest. Thus, even pulses of order 1 volt on 50 ohms could be of significance for a different geometry or for a circuit with a different transfer function. In different configurations the pulsed voltage, or energy, from the same kind of material could be much larger, especially where the vacuum space surrounding the sample contains high electric fields (> 100 V/cm) from any source.

Configurations 7 and 8 need a brief discription. Channel 3 contains a sample of MEP circuit board made of G10 material. This is a nine layer board without any mounted components. Approximately one quarter of the circuit board traces are connected to the pulse detector, the rest of the traces are connected to the instrument ground. The surface of the board is treated with urethane (solithane, tm) conformal coating. A ground plane is conductively glued to most of the back side of the board. Under 200 keV to 1.4 MeV electron beam testing, similar circuits were found to produce 20 to 40 volt peak pulses several nanoseconds wide [2]. Channel 14 contains a similar sample of MEP board, but it is covered with "leaky paint" [2]. Samples with leaky paint were seen to pulse only rarely and with pulses less than 1 volt in ground tests [2].

The leaky paint is an electrically grounded surface coating of semi-insulator with a surface resistivity of the order 10^{10} ohms per square. This level of conductivity prevents surface charging but will not electrically short or load the adjacent circuits.

Signal generators (square pulse) were used to produce known pulsed voltage levels at the sample electrodes and thereby determine the thresholds for pulse detection in each channel. In space on the IDM samples, the pulses are expected to vary in width from less than 1 ns to as much as 20 ns [1]. The pulse detector thresholds vary depending on pulse width and sensitivity setting. Broadly speaking, the thresholds vary from a tenth of a volt to ten volts on the 50 ohm detectors. Much larger pulses would result from the same phenomena occurring on larger samples or on samples with applied voltage on the electrodes [3,4]. The pulses monitored on these samples are assumed to reflect the phenomena of satellite anomalies produced by discharges on wiring, solar array insulators, antenna insulators, circuit boards, feed-thru insulators, and other electrical component insulation.

The size (total energy) of pulses scales linearly with the area of the dielectric which is charged by the radiation [3]. Since the size of IDM samples is small relative to most spacecraft applications, the IDM instrument was set to be sensitive to small pulsed voltages.

CORRELATION OF PULSES WITH SPACE RADIATION

The radiation detectors on CRRES provide a good measurement of the radiation spectra which impinge on the IDM. The spectra of electrons incident on the samples should be calculated by transporting the incident CRRES spectra through the 0.20 mm aluminum cover sheet. Pulse rates would then be correlated with the spectra which directly impact the samples. Since the cover sheet stops all electrons below 150 keV, and passes unimpeded the isotropic electron flux above 1 MeV, the effect of the cover sheet can be ignored for electrons above 1 MeV. The effect on electrons from 150 keV to 1 MeV has not yet been determined. The data presented here is for electrons without correction for the cover plate.

Pulsing first began on one sample on the seventh orbit. The second sample began pulsing on the ninth orbit. The third sample began pulsing on the eleventh orbit. It takes time to get the samples charged up to a near pulsing level.

Figure 2 shows the time history of high energy electron flux during the first 350 orbits. This may be compared to the time history of the pulsing rate to form one of many correlations which are available. The time history of the electron flux from 150 keV to 1.25 MeV has not yet been fully determined so it is not presented. However, the intensity of the electrons in this energy interval usually rises and falls at the same times as does the intensity of the electrons above 1.25 MeV.

Figure 3 shows the time history of pulses from the IDM. As time goes on, the pulse rates from some samples are increasing while others are decreasing.

Comparison of figures 2 and 3 provides an important finding which is in agreement with the accumulation of many laboratory experiments [5], but which may contradict an individual experiment. As an overall average, the pulse rate is crudely proportional to the flux of high energy electrons. However, some samples may pulse profusely at the beginning and pulse sparsely after extended radiation time. Other samples may pulse only after extended irradiation or after extended exposure to vacuum. Two small samples cut from a larger piece may show very different pulsing statistics.

It is tempting to deduce a floor for the electron flux which is required to cause pulsing. Figures 2 and 3 would lead one to think that this can be done. We do not yet do this for two good physical reasons.

 1. We do not yet have good data for the flux of electrons from 150 keV to 1.25 MeV. This component of the energy distribution is probably important for charging of samples. We are now working on this part of the CRRES data base. It requires that we transport the electrons through the 0.2 mm aluminum cover sheet.

 2. The high energy electron fluxes which pass through the samples cause the samples to be more conductive and thereby "leak off charge and lower the peak electric field in the samples." This added conductivity can reduce the probability of pulsing. Using figures 2 and 3 to choose a floor below which one will not have pulsing presumes that this "leak off charge" process is minimized in the IDM samples. No attempt has been made to minimize the leak off process. For example, perhaps a sample mounted on a long narrow insulating post would pulse at much reduced fluxes. Any floor for pulsing deduced from this data would be true only for this set of samples, and can not be extrapolated to other cases.

ORBITAL FLUENCE OF 1.91 MEV ELECTRONS
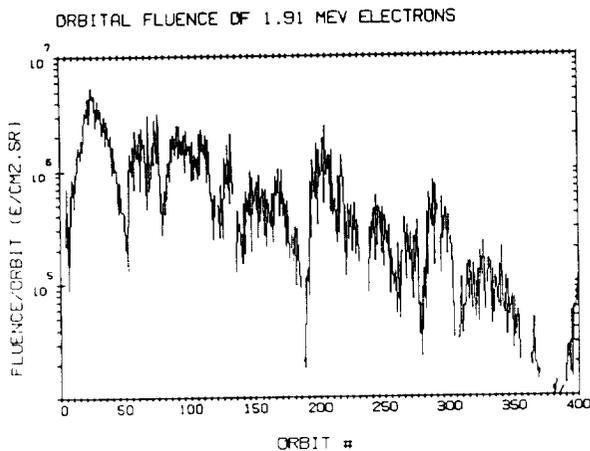
Figure 2. Time History of the Flux Rate From
Electrons in a 400 keV Channel at 1.91 MeV.
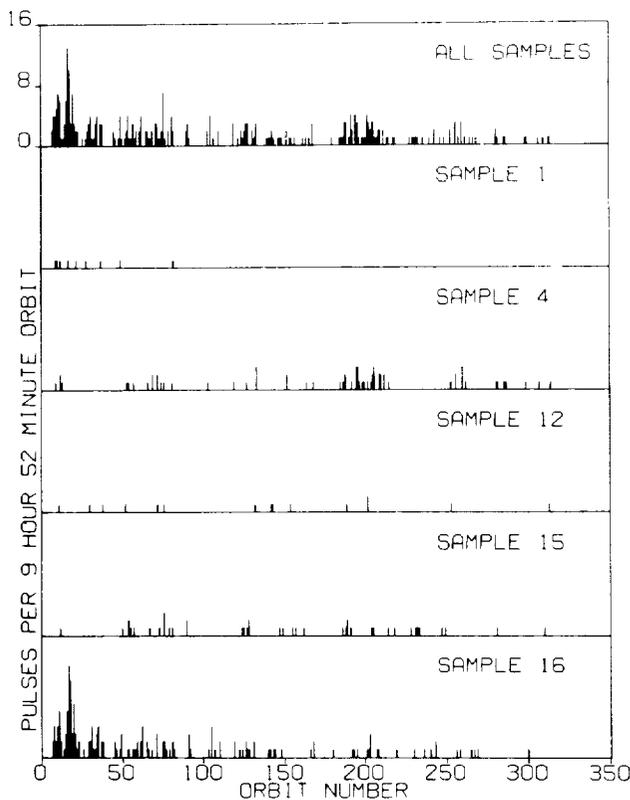
Figure 3. Time History of the Pulse Rates from the
IDM Samples. Most of the samples have never
pulsed. The sample number corresponds to Table I
and is indicated on its graph. These are bar graphs
with one bar for each of the orbits. The height of
the bar indicates the number of pulses during that
orbit.

Nevertheless, we do wish to point out that the
samples' pulse rate decays with time much as the
electron intensity decays as shown by figures 2 and
3. There is a feeling among some in the spacecraft
community that pulsing can be correlated with
outgassing from the samples. The IDM results do
not show that such a correlation is significant. These
samples have been in space a long time. There was
a large increase of the electron flux around orbit 600
and the pulse rate responded nearly in proportion. If
outgassing is an important parameter, then the IDM
samples must have the unusual property of
outgassing at a nearly constant rate since the tenth
orbit, a period of eight months.

Comparison of figure 3 with Table 1 puts a stop to
some common presumptions that were based upon
limited laboratory experiments. For example, one
should not predict the rate of discharging from one
sample based on the rate experienced from a similar
sample; see reference 5 for more on this issue. All
the samples in Table I were seen to pulse in
laboratory tests [1], but only five have pulsed in
space in the first 400 orbits (5 months). It is
interesting that the samples with configurations 3-8
in Table I have not pulsed. Some have predicted
that configurations 3-8 would pulse the most because
of the floating metal or other metal edges on the
samples. The IDM data indicates that one should
not focus on floating metal as an issue of any
significance for radiation-induced spacecraft
anomalies. For more ideas concerning the variability
of pulsing phenomena from apparently similar
samples, see references 5-8. The floating metal
issue is commented in reference 8 at the bottom of
page 228.

The IDM is beginning to accumulate enough data
that we may speculate on the causes of the pulsing.
Figure 4 is developed to help us do that. In this
figure we present the time history of the CRRES
IDM during one average orbit which is a composite
of the first 350 orbits. We begin the orbit at apogee
where the electron flux is low. The curve is a trace
of the average electron flux experienced by the IDM
during the 350 orbits. The electron flux rises as the
satellite moves from apogee at the left end of the
graph. On average, the satellite experienced the
peak electron fluxes (at an L shell of 3.5 to 4.5)
about 13000 seconds after apogee. The electron
fluxes diminish around perigee and then peak up
again at 24000 seconds after apogee. The satellite
attains apogee again at roughly 36000 seconds.

782
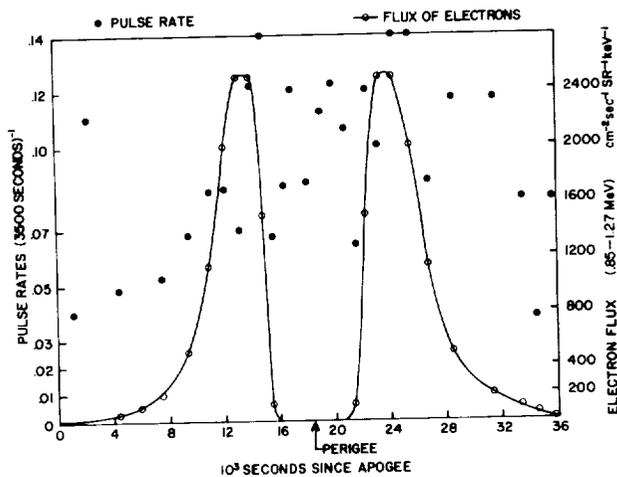
Figure 4. An Average Orbit Time History of the High Energy Electron Flux and IDM Pulse Rates Starting and Ending at Apogee. The solid dots are the measured IDM pulse rates. The flux data is for one electron energy channel, 0.85 to 1.27 MeV. Inclusion of all electron energies would not change the shape of this curve in any significant way.

The pulse rate is shown as a histogram on figure 4. We collected data in 1/2 L shell bins so that there are 28 separate "bins" containing pulse rate data on this graph. Note that CRRES apogee varies, typically from L=7 to L=9, and everything above L=7.5 is accumulated as one bin; this is a reasonable thing to do because the flux is not contributing much here. Similarly, the data from "L=0" to L=1.5 is lumped together. With so many bins and only 360 pulses, the statistics of the scatter data is not good.

The pulse rate rises to a maximum after passing through the electron belt on its outward trajectory at around 25000 seconds. Thereafter the pulse rate drops almost in half after the satellite passes through apogee and approaches the electron belt on its inward pass. This decay of pulse rate at high altitudes while near apogee is in reasonable agreement with laboratory experience as reported in reference 5, page 347, where samples made with fiberglass pulsed for days after the irradiation ended. Based on figure 4, very loosly speaking, one could say that the pulse rate, for average IDM pulses, decays in half after about 5 hours without radiation. The time spent at perigee without electron irradiation, being only 2 hours, is not long enough to see any decay in the data.

Figure 4 presents an important fact that must be remembered when interpreting IDM (or other spacecraft anomaly) pulse data. Pulses do not occur because of the instantaneous radiation flux intensity. Pulses result from the accumulation of charge in the insulator which produces high electric fields. The high electric fields, acting alone, do not produce the pulses either. Instead, at the magnitude of electric fields expected [8] in these samples, typical insulators occasionally and spontaneously produce pulses called prebreakdown events. The high voltage insulator literature has much information on this phenomena, yet it cannot be predicted apriori with certainty. The electric fields in these samples decay very slowly so that pulsing can continue for hours or days after radiation ceases. There should be no surprise in this result, those who work with insulators know of the phenomena.

FUTURE WORK

Having determined that the pulsing samples are responding to an orbit averaged flux, and not to the instantaneous flux of electrons, we can move on to a search for more correlations. We have begun such a search but the work is tedious. We plan to correlate the pulse rate with the hardness of the electron spectrum. There is an intriguing hint in the raw data that one pulse on a sample causes another pulse to be more likely in the immediate future. Leon Levy has noted this in laboratory experiments and Levy and Frederickson have posited an explanation for it (Presented at URSI, Fall 1990). Based on this model one would be tempted to look for enhanced rates of pulsing due to passage through the proton belts, but such correlations would need to be rationalized with the known effects created by the very large range of proton energies in those belts. At the end of March 1991 the electron belts were enhanced a factor of 50 to 100 and produced extreme levels of pulsing on the IDM. The total pulse count more than doubled in a week, and it was the same samples as before. This new data improves our statistics and will make future correlations even more meaningful.

SUMMARY

PULSE RATES APPEAR TO SCALE APPROXIMATELY LINEARLY WITH ELECTRON FLUENCE PER ORBIT.

SAMPLES WITH FLOATING METAL ARE RARELY PULSING, THE SIMPLE GEOMETRY IS PULSING THE MOST.

APPROXIMATELY 600 PULSES HAVE BEEN SEEN IN NINE MONTHS SINCE LAUNCH.

TOTAL SAMPLE AREA IS 300 CM^2. SCALING TO FULL SATELLITE AREA IMPLIES THAT MANY PULSES PER SATELLITE MAY OCCUR.

PULSES OCCUR EVEN HOURS AFTER PASSING THROUGH THE ELECTRON BELTS. PULSE RATES DECAY WITH APPROX 10 HOUR TIME CONSTANT OUTSIDE THE ELECTRON BELTS.

THESE RESULTS ARE FOR SAMPLES WITH NO BIAS. APPLIED BIAS COULD ENHANCE PULSE RATES OR PULSE MAGNITUDES.

THERE SEEMS TO BE NO CORRELATION OF PULSING WITH OUTGASSING.

THERE HAVE BEEN PERIODS OF LOW ELECTRON FLUX WHERE THE PULSE RATE APPROACHES ZERO, ABOUT 1 PULSE PER WEEK. MAXIMUM RATES ARE ROUGHLY 5 PER HOUR.

ACKNOWLEDGEMENT

REFERENCES

1. P. G. Coakley, M. J. Treadaway and P. A. Robinson, Jr., "Low Flux Laboratory Test of The Internal Discharge Monitor (IDM) Experiment Intended for CRRES," IEEE Trans. Nuc. Sci. 32, No. 6, (Dec 85) 4066-72.

2. A. R. Frederickson, et. al., "Leaky Insulating Paint for Preventing Discharge Anomalies on Circuit Boards," IEEE Trans. Nuc. Sci. 36, No. 6, (Dec 89) 2405-10.

3. K. G. Balmain and G. R. Dubois, "Surface Discharges on Teflon, Mylar and Kapton," IEEE Trans. Nuc. Sci. 26, 5146, (1979).

4. A. R. Frederickson, L. Levy and C. L. Enloe, "Radiation-induced Electrical Discharges in Complex Structures," to be published in IEEE Trans. Elec. Insul.

5. A. R. Frederickson, "Electrical Discharge Pulses in Irradiated Solid Dielectrics in Space," IEEE Trans. Elec. Insul. 18, 3, 377-49, (1983).

6. P. Coakley, M. Treadway, N. Wild, and B. Kitterer, "Discharge Characteristics of Dielectric Materials Examined in Mono-, Dual-, and Spectral Energy Electron Charging Environments," in Spacecraft Environmental Interactions Technology 1983, NASA CP-2359 or AFGL TR-85-0018, 511-24.
(NTIS #N85-22470)

7. A. R. Frederickson, "Discharge Pulse Phenomenology," ibid, 483-510.

8. A. R. Frederickson, "Partial Discharge Phenomena in Space Applications," in Fourth European Symposium on Spacecraft Materials in Space Environment, 221-31, ONERA-CERT, Toulouse, France, Sept. 1988.

ENVIRONMENTAL INTERACTIONS IN SPACE EXPLORATION:

ENVIRONMENTAL INTERACTIONS WORKING GROUP

Joseph C. Kolecki
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135

G. Barry Hillard
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135

INTRODUCTION

With the advent of the Space Exploration Initiative as mandated by the President of the United States in his twentieth anniversary speech commemorating the first Moon landing, the possibility of designing and using new space systems on scales not heretofore attempted opens before us exciting new challenges in systems design and space science. The environments addressed by the Space Exploration Initiative include the surfaces of the moon and Mars, as well as the varied plasma and field environments which will be encountered by humans and cargo en-route to these destinations from the surface of the Earth, or from low Earth orbit. It is fully expected that systems designers will have to take into account the environmental characteristics for which they are designing. They will need to understand the interaction mechanisms to which their systems will be subject and be able to model them from the earliest conceptual design stages through to design completion.

Planetary and space environments are interactive with man-made systems in a variety of ways, as experience has already shown. Lunar dust physics is dominated by electrostatic forces. The martian surface environment differs from that of the moon in that it consists both of sand and dust, and a dust-laden, low pressure atmosphere in which Paschen electrical breakdown can occur at voltages ranging from a few hundred volts upward. The space environment is dominated by plasmas and fields with a wide range of energies and other physical characteristics. In addition to these natural environments, humans will produce local environments of their own which will modify what is already there and have to be dealt with on their own terms. The characteristics of system/environmental interactions depends on both the characteristics of the local environment, and those of the system. Different system scale sizes lead to different types of interactions and require different mathematical models to enable understanding, evaluation, and prediction.

To the end of understanding environmental interactions for the Exploration Initiative, an Environmental Interactions Working Group has been established as part of the Engineering Requirements Subgroup. This working group is chaired by Joe Kolecki of NASA Lewis and draws its membership from a broad scope of disciplines. The current paper describes the working group and gives an update on its current activities. The working group charter and operation are reviewed, background information on the environmental interactions and their characteristics as we understand them is offered, and the current status of the group's activities are presented along with anticipations for the future.

ENVIRONMENTAL INTERACTIONS WORKING GROUP

Space system/environmental interactions comprise a set of phenomena which occur when a system is placed into an environment whose local characteristics are such that the system and the environment are able to exchange "information" in some way and thereby influence or modify each other. Put another way, one can imagine the various environments of concern to be characterized each by some parameter set (like an Earth Standard Atmosphere model), and the various systems likewise. An environmental interaction is what occurs when one takes a system from shelf "A," an environment from shelf "B," places them into a sealed bag and shakes well. It is the objective of an environmental interactions team to classify, prioritize, and model interactions, to advocate and/or conduct all necessary experiments whether in space or on the laboratory, and to place into the hands of systems designers the best possible user friendly codes and/or design guideline documents possible. This will enable those designers to optimize their systems to the environments in which they must function, from the earliest design stages through to completion and use.

This type of work is already being carried out in a number of discipline areas, some of which have been ongoing for a considerable time. Some of these areas include radiation effects, having to do with particles and fields; collision effects, having to do with micrometeoroids and orbital debris; effects connected with dusty plasmas, like those found on the moon or in the vicinity of comet

nuclei; plasma and spacecraft charging effects, chemical processes, Martian aeolian processes, and temperature cycling effects. Systems which have historically been impacted by interactions with their local environments include the "big" missions systems like those associated with Apollo, Viking, and the planetary and cometary probes of the past few decades, geostationary satellites, which displayed anomalous behavior most readily accounted for by charge/discharge phenomena, the Space Shuttle, which has carried various payloads to determine environmental parameters at LEO (some of which were used in development of computer models) and the Space Station, whose size and complexity has led to a reconsideration of environmental interactions in LEO. The evolution in our knowledge of environmental interactions has been driven by the growth and variety of space infrastructure elements placed onto Earth orbit and elsewhere. Systems of increasing size and complexity, and placement of those systems into active new environmental regimes has repeatedly necessitated the development of new interactions knowledge and system design guidelines. The President's Exploration Initiative is orders of magnitude larger than anything attempted in space to this point in time. Therefore, our understanding of environmental interactions must again grow to meet the challenge of the future.

The environments of concern to the Exploration Initiative include at least the surface environments of the Moon and Mars, the martian atmosphere, and the space environments through which humans and cargo will travel to reach these destinations. Least known among these environments are the martian surface and atmosphere. Interactions to be anticipated when systems are placed on the martian surface include Paschen breakdown from moderate to high voltage surfaces, charge/discharge phenomena associated with electrostatically charged dust, modification of discharge onset characteristics depending on wind and wind-borne dust conditions, and charge/discharge phenomena during descent and ascent of man-made vehicles. Dust deposition will modify surface optical, electrical, and thermal properties, and wind borne dust will produce mechanical erosion over long periods of time. Chemical effects are also to be anticipated with active species in the martian atmosphere and soil. Additional thoughts on martian environmental effects are included in the appendix to this abstract.

Better known is the lunar environment in which dust dynamics are driven by electrical forces, and result in copious contamination of external and internal surfaces and volumes. Since the moon has no atmosphere, radiation and plasma effects are also important lunar surface considerations. Long term human occupation of the Moon and Mars necessitates development of systems in which interactions with the environment are understood in the earliest design phases, and taken into account throughout the design process. Similar arguments exist for translunar and interplanetary space through which humans

and cargo must pass en-route to lunar and Martian destinations, as well as to low Earth orbit which may be the site of semi-permanent human occupation in an Earth orbiting Space Station. The size and complexity of the Space Station have already been cited as drivers for developing new environmental interactions knowledge, and the issues which have arisen (with such design parameters as Space Station structure to solar array grounding) must serve as fair warning of the need for constantly updating our interactions models and/or developing and experimentally validating new ones, as necessity determines.

To date, environmental interactions work has been carried out in several specific areas and completed to varying degrees. But this work has not been carried out in the "Integrated Program" fashion necessary to the scope and magnitude of a Space Exploration Initiative. The respective disciplines have evolved semi-autonomously, that is, to a large extent independently of each other. Cross fertilization of ideas has occurred only as needed, as when the question of a plasma interaction arose around the issue of a solid object like a meteoroid impacting or penetrating a spacecraft bumper shield and generating a hot transient gas as a by-product. No forum yet exists in which system/plasma interactions are discussed on a large scale, with cross-disciplinary exchange of ideas, and development of integrated program elements which both insures minimal overlap of funded efforts and fills all of the programmatic "holes."

An integrated, coordinated approach is required for future work in this area, allowing programs to develop with mutual knowledge of each other and knowledge of the mission and system definition activities already under way. To that end, a working group has been established to address issues and make recommendations, defining agency-wide areas of need for predicting and evaluating interactions, establishing what information is already available, and formulating and implementing a program to develop and experimentally validate models for use by system design engineers. As a point of ingress into a potentially infinite area of study, the Environmental Interactions Working Group will begin by establishing requirements for robotic missions to Mars. The group will establish a data base of currently available information on the Martian surface environment, and understand in broad terms the expected operating parameters of systems intended for use on the martian surface. It will then define and prioritize martian surface interactions, and establish those parameter sets most critical to developing interactions models. The parameter sets will be compared against existing data, and used to define instruments, sensors, and/or experiments to be carried to the martian surface as part of the robotic precursor "package." These items will be documented in the form of a set of recommendations, and delivered to the Engineering Requirements Working Group being established at the Johnson Space Flight Center. The working group will

expand its activities to embrace the broader charter described above as appropriate. It will meet by videoconference approximately every six weeks, and more frequently on an as-needed basis.

APPENDIX: OUTLINE OF PRELIMINARY THOUGHTS ON MARS SURFACE INTERACTIONS

The following outline represents preliminary thoughts on electrical system/environment interactions on the planet Mars. These types of interactions should be of interest to system designers at every stage of system design.

I. ATMOSPHERIC INTERACTIONS
   A. Paschen electrical breakdown from high voltage surfaces: (The 7-9 Torr atmospheric pressure at the martian surface is ideal for Paschen breakdown over mm to cm distances at a few hundred volts, and cm to m at a few kilovolts)
      1. breakdown to other spacecraft surfaces
      2. breakdown to air
      3. breakdown to ground
      4. source of system power loss
      5. source of electromagnetic noise
   B. Sparks, arcs, and other sustained or transient discharge phenomena
      1. electrical power loss
      2. sputter erosion of surfaces
      3. transport and redeposition of materials (contamination)
   C. Electrical discharge phenomena during descent and ascent (unknown)
      1. electrical breakdown of low pressure atmosphere and local gases due to thruster firing
      2. differential charging of surfaces during thruster firings
      3. analogous to discharge events seen during terrestrial launches

II. MARTIAN IONOSPHERE/SPACECRAFT INTERACTIONS IN LMO (100-300KM)
   A. Plasma phenomena are functions of system voltage, system configuration, and local particle populations
   B. Max neutral number density = 10^9 and 10^{12} cm^{-3} for monatomic oxygen and CO_2 respectively (compare 10^9 cm^{-3} for monatomic oxygen in LEO)
   C. Max electron density = 10^4 to 10^5 cm^{-3} (compare 10^5 to 10^6 cm^{-3} in LEO)
   D. Max ion density = 10^4 to 10^5 cm^{-3} for CO_2^+ and O_2^+ respectively (compare 10^4 to 10^6 cm^{-3} for N^+ and O^+ respectively in LEO)
      1. PIX II showed evidence of arcing and parasitic current collection at 10^4 cm^{-3} electron number density
      2. PIX II operating voltages = $+/- 750$ volts

E. Concerns include spacecraft surface charging, sputtering from charged surfaces, spacecraft floating potentials, oxygen degradation of materials, and spacecraft system grounding

III. DUST RELATED ELECTRICAL INTERACTIONS
   A. Surface dust
      1. charging due to solar ultraviolet
      2. charging due to induced dipole effects from exposed high voltage surfaces
      3. deposition on surfaces due to electrostatic attraction
      4. modification of surface thermal, optical, and dielectric properties
      5. contamination of clean volumes
   B. Wind borne dust
      1. charging due to frictional mechanisms
      2. charging due to differential settling after dust storms
      3. deposition of electrically charged dust on surfaces
      4. modification of surface thermal, optical, and dielectric properties
      5. contamination of clean volumes

IV. INTERACTION MODELS: WHAT HAPPENS WHEN YOU ADD A SYSTEM TO AN ENVIRONMENT AND "SHAKE WELL?" HOW DOES ONE CHARACTERIZE THE RESULTING INTERACTION AND USE THAT KNOWLEDGE TO PRODUCE THE BEST DESIGN? SOME AREAS FOR CONSIDERATION/ DEVELOPMENT ARE THE FOLLOWING:
   A. Identify relevant physical mechanisms/equations
   B. Produce, and (where necessary) experimentally verify mathematical models
      1. identify existing interaction models where applicable
      2. identify and direct models currently under development in related areas
      3. identify areas for which model development needs to be initiated (SEI specific models, or working modules for use with existing models)
   C. Understand user needs and establish appropriate user interfaces
      1. system designers are the most likely users
      2. interaction models are a bridge between standard environment models and engineers who wish to understand how their new system will play in a given environment
   D. Establish appropriate input/output formats
   E. Identify/perform relevant laboratory/space tests and analyses
   F. Deliver user-friendly software with appropriate interfaces

787

Medium Resolution Spectra of the Shuttle Glow in the Visible Region of the Spectrum

R. A. Viereck, E. Murad, and C. P. Pike
Phillips Laboratory, Spacecraft Interactions Branch

S. B. Mende, and G. R. Swenson
Lockheed Missile and Space Co.

F. L. Culbertson and B. C. Springer
NASA, Johnson Space Center

Recent spectral measurements of the visible shuttle glow (λ = 400 - 800 nm) at medium resolution (1 nm) reveal the same featureless continuum with a maximum near 680 nm that has been reported previously. This is also in good agreement with recent laboratory experiments that attribute the glow to the emission from NO_2^* formed by the recombination of O + NO. The data that will be presented were taken from the aft flight deck with a hand-held spectrograph and from the shuttle bay with a low-light-level television camera. Shuttle glow images and spectra will be presented and compared with laboratory data and theory.

A Comparison of Shuttle Vernier Engine Plume Contamination with CONTAM 3.4 Code Predictions

Carl R. Maag, Thomas M. Jones* and Shankar M. Rao*
Science Applications International Corporation
2605 E. Foothill Blvd, Suite A
Glendora, CA 91740
818-335-6888

and

W. Kelly Linder, Captain
United States Air Force, Space Systems Division
NASA/Lyndon B. Johnson Space Center
Houston, Texas 77058

In 1985, using the CONTAM 3.2 code, it was predicted that the shuttle Primary Reaction Control System (PRCS) and Vernier Reaction Control System (VRCS) engines could be potential contamination sources to sensitive surfaces located within the shuttle payload bay. Spaceflight test data on these engines is quite limited. Shuttle mission STS-32, the Long Duration Exposure Facility (LDEF) retrieval mission, was instrumented with an experiment that has provided the design engineer with evidence that contaminant species from the VRCS engines can enter the payload bay.

Before proximity operations began with the LDEF, the orbiter executed a maneuver that involved firing a VRCS engine. Within a second of the thruster firing, a significant mass deposition was recorded by Temperature-controlled Quartz Crystal Microbalances mounted within the experiment. Postflight chemical analysis of the sensing crystals and surrounding areas, showed organic acid salts and complex nitrates present.

More recently, the most recent version of the analysis code, CONTAM 3.4, has re-examined the contamination potential of these engines. This paper will present a comparison of the most recent analytical assessment of the VRCS engine along with empirical data from the STS-32 mission. A description of the experiment will also be included.

* 1875 Century Park East, Suite 2190
  Los Angeles, California 90067

I

AUTHOR INDEX

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE February 1992	3. REPORT TYPE AND DATES COVERED Conference Publication

4. TITLE AND SUBTITLE
Fifth Annual Workshop on Space Operations Applications and Research (SOAR '91)

5. FUNDING NUMBERS

6. AUTHOR(S)
Kumar Krishen, Editor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Lyndon B. Johnson Space Center
Houston, Texas 77058

8. PERFORMING ORGANIZATION REPORT NUMBER
S-650

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, D.C. 20546
U.S. Air Force, Washington, D.C. 23304
University of Houston - Clear Lake, Houston, Texas 77058

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA CP-3127, Vol. II

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified - Unlimited

Subject Category: 59

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This document contains papers presented at the Space Operations, Applications and Research Symposium, hosted by NASA Johnson Space Center (JSC) and held at JSC in Houston, Texas, on July 9-11, 1991. More than 110 papers were presented at this Symposium, sponsored by the U.S. Air Force Phillips Laboratory, the University of Houston-Clear Lake, and NASA JSC. The technical areas covered were Intelligent Systems, Automation and Robotics, Human Factors and Life Sciences, and Environmental Interactions. The U.S. Air Force and NASA programmatic overviews and panel discussions were also held in each technical area. These proceedings, along with the comments and suggestions made by the panelists and keynote speakers, will be used in assessing the progress made in joint USAF/NASA projects and activities. Furthermore, future collaborative/joint programs will also be identified. The SOAR '91 Symposium is the responsibility of the Space Operations Technology Subcommittee (SOTS) of the USAF/NASA Space Technology Interdependency Group (STIG). The Symposium proceedings includes papers covering various disciplines presented by experts from NASA, the Air Force, universities, and industry.

14. SUBJECT TERMS
Space, operations, automation, robotics, life sciences, medical protocols, knowledge acquisition, expert systems, radiation, biomedical, debris, space shuttle, Space Station Freedom, lunar outposts, Mars missions.

15. NUMBER OF PAGES
392

16. PRICE CODE
A17

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT