

A Methodology for Assessing the Military Benefits of Science and Technology Investments

By Albert Sciarretta, Richard Chait, Joseph Mait, and Jordan Willcox

**Center for Technology and National Security Policy
National Defense University**

September 2008

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE SEP 2008		2. REPORT TYPE		3. DATES COVERED 00-00-2008 to 00-00-2008	
4. TITLE AND SUBTITLE A Methodology for Assessing the Military Benefits of Science and Technology Investments				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Defense University, Center for Technology and National Security Policy, 300 5th Avenue SW, Washington, DC, 20319				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 59	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

The views expressed in this report are those of the authors and do not reflect the official policy or positions of the National Defense University, the Department of Defense, or the U.S. Government. All information and sources for this paper were drawn from unclassified materials.

Albert Sciarretta is a Senior Research Fellow at the Center for Technology and National Security Policy (CTNSP). He is a retired Army officer who is considered an expert in assessing advanced military technologies, designing operational experiments, and developing investment strategies. He received two M.S. degrees—Operations Research and Mechanical Engineering—from Stanford University and a B.S. degree from the U.S. Military Academy.

Richard Chait is a Distinguished Research Fellow at CTNSP. He was previously the Chief Scientist, Army Materiel Command, and Director, Army Research and Laboratory Management. Dr. Chait received his Ph.D. from Syracuse University and a B.S. degree from Rensselaer Polytechnic Institute.

Dr. Joseph N. Mait is a Senior Technical Researcher at the U.S. Army Research Laboratory. He is the Cooperative Agreement Manager of the Collaborative Technology Alliance on Micro-Autonomous Systems and Technology. Dr. Mait received his Ph.D. and M.S. in Electrical Engineering from Georgia Tech, B.S. degree in electrical engineering from the University of Virginia.

Jordan Willcox is a Research Assistant at CTNSP. He received his B.A. in International Relations from Connecticut College and is pursuing an M.A. in Security Studies from Georgetown University. Mr. Willcox gained experience in policy analysis, research support, and database design at the Nixon Center and the Organization for American States.

Acknowledgements. The authors wish to acknowledge Dr. Thomas Killion, Army Science and Technology Executive; Dr. Hans Binnendijk, Director of CTNSP; and Dr. Michael Baranick, Senior Research Fellow at CTNSP, for their support of this effort. It should be noted that Dr. Killion also co-authored the unpublished paper, “Measuring Return on investment for Army Basic Research,” provided as appendix C. Finally, we thank our colleagues Dr. John Lyons and Dr. Charles Barry for contributing information on the importance of globalization and joint capabilities, respectively.

Defense & Technology Papers are published by the National Defense University Center for Technology and National Security Policy, Fort Lesley J. McNair, Washington, DC. CTNSP publications are available online at <http://www.ndu.edu/ctnsp/publications.html>.

Contents

Chapter 1. Introduction	1
Chapter 2. Difficulty with Assessing a Return-on-Investment Effect in the Military	6
Commercial ROI	6
The Problem for the Army	7
Recent Efforts	8
Solving The Problem is Not Impossible	9
Chapter 3. Cost-Benefit Analysis	10
Categories of Military Benefits	11
Assessment of Categories of Military Benefits	13
Completing the Cost-Benefit Analysis	18
Chapter 4. Composite Index	19
Chapter 5. Probability of Achieving the Anticipated Benefit	21
Components of the Probability	21
Calculating the Probability	24
Chapter 6. Project Portfolio Management Tools	25
PALMA	27
Other Evaluated Tools	29
ExpertChoice®	29
Equity	30
Balanced Scorecard	31
Chapter 7. Demonstrating the Methodology	32
Chapter 8. Summary and Concluding Remarks	34
Appendix A. Assessment of UK Approach to Quantifying Benefits of R&D	36
Appendix B. Effects of Globalization on Army Research and Development	38
Appendix C. “Measuring Return on Investment for Army Basic Research”	40
Appendix D. Importance of Joint Capabilities in Cost-Benefit Analysis Efforts	50
Appendix E. Development of Metrics	51
Determining Metrics for an S&T Program	51
Technical Approach for Developing Metrics	52
Example of a Documented MOE	53
Appendix F. Acronyms	54

Chapter 1. Introduction

This paper originated with the concerns of the Deputy Assistant Secretary of the Army for Research and Technology (DASA(R&T)), whose duty is to assess the Army's science and technology (S&T) program.¹ The Deputy Assistant Secretary has aggressively sought innovative ideas for measuring the impact of the Army's S&T on the future fighting force. Recently, the National Defense University's Center for Technology and National Security Policy (CTNSP) conducted a full review of the past contributions of Army laboratories to today's military capabilities.² Aware of that review, the Deputy Assistant Secretary asked CTNSP to develop approaches for measuring the benefits of today's S&T investments on the future military.

Management of the S&T process to obtain the best "bang for the buck" has changed considerably through the years. As documented by Miller and Morse,³ the research and development (R&D) process as we know it can be traced back to Thomas Edison's establishment of a research laboratory in New Jersey during the late 1800s. Other laboratories soon followed, so that by the early 1900s corporations such as General Electric, Eastman Kodak, Bell Telephone, and Dupont all had active R&D facilities. These efforts were highly successful, providing innovations that led to many useful and profitable products.

The importance of research to a corporation's future success is also highlighted by Buder⁴ in his examination of the role of corporate research in some of the world's top information technology companies. He concludes that corporate research is more connected today to its original purpose of advancing company goals than at any time in the past half century.

In these early R&D laboratories, management was provided by scientists who selected and conducted the research. Innovations from the R&D in these laboratories played significant roles in the outcomes of both World War I and World War II. Industrial, academic, and government in-house laboratories produced technologies that led to new allied air, ground, and sea warfare capabilities. These technologies included radar and sonar, synthetic rubber, proximity fuses, the ENIAC,⁵ and the atomic bomb.

In the years following World War II, corporate R&D groups continued to develop new technology in pursuit of market opportunities. This technology push contributed to rapid expansion of the economy, as well as increases in both R&D activity and associated costs. As a

¹ S&T is a subset of research and development (R&D). S&T includes basic research, applied research, and advanced technology development. R&D includes S&T as well as advanced component development and prototypes.

² John Lyons, Richard Chait, and Duncan Long, "Critical Technology Events in the Development of Selected Army Weapons Systems," *Defense & Technology Paper 35* (Washington, DC: Center for Technology and National Security Policy, September 2006).

³ William L. Miller and Langdon Morris, *4th Generation R&D; Managing Knowledge, Technology, and Innovation* (New York: John Wiley and Sons, Inc., 1999).

⁴ Robert Buder, *Engines of Tomorrow, How the World's Best Companies Are Using Their Research Labs to Win the Future* (New York: Simon and Schuster, 2000).

⁵ Electronic Numerical Integrator and Computer (ENIAC) was the first high-speed, purely electronic, Turing-complete, digital computer capable of being reprogrammed to solve a full range of computing problems. ENIAC was designed and built to calculate artillery firing tables for the U.S. Army Ballistic Research Laboratory.

result, these costs became more visible in a corporation's overall financial picture. In response, some corporations developed methodologies, such as return on investment (ROI) analysis, to analyze R&D costs and potential benefits. Many of these methodologies were modeled after procedures for monitoring capital investment.

This closer control of R&D expenditures led to a sharper focus on customer needs. Technological development efforts were increasingly directed by assessments of these needs. Approaches such as strategic planning and technology roadmaps began to be used as part of the assessment and development process. In the 1990s, companies began to use a business process focused on the entire innovation process, rather than just on R&D. In essence, this process, known as continuous innovation, involved collaboration of effort (or teaming) both within the entire organization and with suppliers and customers.

During this time period, similar teaming arrangements also proved to be beneficial to the Department of Defense (DOD) as it fielded advanced weapon systems. For example, a recent in-depth study of fielded Army weapon systems,⁶ revealed the importance of teaming among in-house laboratories, industry, and academia. In this study, the origin of critical technology events—best understood as ideas, approaches, concepts, models, and analyses that had a major impact on the development of a specific weapon system—was described for four weapon systems: the Abrams tank, the Apache helicopter, and Stinger and Javelin missiles. One of the key findings of these retrospective studies was the importance of the ability of the in-house laboratories, industry, and academia to work smoothly across organizational lines. The attitude of senior managers concerning teamwork with other stakeholders strongly influenced how those at the R&D bench conducted their work. With this in mind, some of the Services have created large, long-term research centers that put more emphasis on certain key technology areas and promote participation by the private sector and academia. For example, the Army has funded the creation of the Institute for Soldier Nanotechnologies at the Massachusetts Institute of Technology (MIT), Cambridge, MA, where research will be conducted by both universities and private corporations.

Retrospective studies such as those just described provide valuable information on the origin of new technologies and the role of in-house laboratories, industry, academia, and international R&D, but do not provide quantitative information on the role of R&D in determining the quality of military systems.

Other techniques in the public literature for assessing commercial ROI were also judged as inadequate to fully solve the military problems addressed by our study.

A significant effort to quantify the value of S&T efforts was made by Dr. Edwin Mansfield of the National Institute for Standards and Technology in the 1970s. His paper on a cost-benefit methodology for S&T programs, did much to quantify the ROI for publicly funded S&T.⁷

⁶ John Lyons, Richard Chait, and Duncan Long, "Critical Technology Events in the Development of Selected Army Weapons Systems," *Defense & Technology Paper 35* (Washington, DC: Center for Technology and National Security Policy, September 2006).

⁷ Edwin Mansfield, *Toward a Standard Benefit-Cost Methodology for Publicly Funded Science and Technology Programs*, NIST IR 7319.

However, his focus on monetary gains does not fit well in assessing military S&T, as will be discussed in detail in chapter 2. Another noteworthy attempt to confront this analytical challenge was made in a recent study by Middleton, et al.⁸ The study was based on an examination of weapon systems from ten countries, including the United States, UK, Russia, France, and Japan. In all, some 69 military systems with the following military service distribution were judged for quality: Army (25 systems), Navy (23 systems), and Air Force (21 systems). The Middleton approach is valuable for giving a rough idea of national technological advancement, but it is not quantitative enough nor sufficiently detailed to evaluate the causal effect of S&T on specific military capabilities, as needed by the DASA(R&T) to assess Army S&T. The Middleton approach was a low-resolution, high-level, highly subjective analysis of overall military capability for a given national level investment of R&D funds. (A more detailed discussion of the Middleton effort is presented in appendix A.)

When assessing the benefits of S&T, the tools used must recognize that unlike commercial items, where profit and market share are important objectives, other defining factors—such as military capability enhancements—are key to weapon system quality. As an example of the influence of Army R&D expenditures upon weapon system capability, prior studies have documented how investments in energetic materials and penetration mechanics in the early 1980s led to the development of the M829A1 projectile. The effectiveness of the M829A1 was one of the factors in the overmatching tank-killing capability developed during Operation *Desert Storm*. In addition to military capability, other critical factors influence the benefits of R&D. This paper will describe these factors in detail and incorporate them within a methodology that provides a benefit-to-cost ratio for technological development.

It should be noted that the methodology will need to be implemented in a manner sensitive to what components are appropriate for the technology under evaluation. The design of the implementation will be dependent on the objective use of the technology being assessed, the ability to quantify future benefits of the technology, the availability of programmatic information, supporting resources (e.g., tools and funding), and time available for assessment. For example, if the objective use of the technology is to improve manufacturing processes, then it would be inappropriate to assess its impact on measures of effectiveness for combat operations.

The methodology to be discussed in this paper has four major components:

1. “ROI-like” assessments focused on typical capability, effectiveness, and performance parameters as well as other areas of benefit (programmatic, life-cycle costs, manufacturability, etc.). Modeling and simulation tools may be used to provide predictive measures of capability, effectiveness, and performance.
2. Development of a “composite ROI index” that would group many similar terms and can be plotted against funding. This index would include the sum of multiple terms, with each term having a weight “W,” a requirement “R” (e.g., objective flight duration), and a measure of performance “P” (e.g., actual flight duration). Element “i” of the index would be $W_i(P_i/R_i)$. The

⁸Andrew Middleton, Steven Bowns, Keith Hartley, and James Reid, “The Effect of Defence R&D on Military Equipment Quality,” *Defense and Peace Economics*, Volume 17, Number 2 (April 2006), 117–139.

composite index may be most appropriately used when dealing with numerous parameters identified in 1 above.

3. Calculation of probability of achieving the measured ROI. This would take into account probabilities of technical success, transition opportunity, technology insertion, adequate funding, and ability to maintain a capability advantage.

4. Utilization of a project portfolio management (PPM) tool to select projects (identified as most promising in the analysis above) for funding based on available budgets, objectives, etc. For this paper, a tool currently being used in the Office of the DASA(R&T), as well as four other PPM tools—PALMA (Portfolio AnaLysis MACHine), Balanced Scorecard, Equity, and Expert Choice—were reviewed, with PALMA being selected as having the most relevance for a follow-on demonstration of the methodology.

The methodology will also incorporate the impact of two growing trends that are already having a significant impact on R&D. One is globalization, which is summarized in appendix B, but in more detail in a separate CTNSP publication.⁹ Globalization is a two-edged sword. On one hand, it affords opportunities for collaboration that could accelerate progress on overcoming mankind's technical limitations. On the other hand, it quickly disseminates technical information, thus significantly reducing the amount of time that a country may hold a technological edge. The methodology attempts to address both the benefits and the problems associated with globalization.

The second trend is internal to the U.S. military—the desire to transform the military into a joint operating force that exploits netcentric capabilities. To support this transformation to “jointness,” netcentric capabilities must be developed that enhance the interoperability of activities, operations, organizations, etc., in which elements of two or more Services participate. Not all technologies need to support this netcentric transformation. For example, developing obstacle avoidance technologies for micro-unmanned aerial vehicles (micro-UAVs) does not support joint operations. However, jointness would be enhanced with a capability to transmit images from Army-operated micro-UAVs to an Air Force close air support mission aircraft. Thus, in assessing R&D, it may be appropriate to give bonus points for jointness, no points for a technology that can play no role in joint operations, and perhaps negative points for a technology that could help support joint operations but is not designed to do so.

In a future paper, we will demonstrate the usefulness of the methodology described herein by applying selected components in a *use case analysis*¹⁰ involving autonomous systems—a technology area we believe will have high payoff for DOD.

In this paper, we begin by first addressing the difficulty the military has had in utilizing ROI techniques for assessing its S&T program. This is followed by discussions of assessing anticipated benefits of S&T investments with a proposed methodology in terms of an ROI-like

⁹ John Lyons, “Army R&D Collaboration and The Role of Globalization In Research,” *Defense & Technology Paper 51* (Washington, DC: Center For Technology and National Security Policy, July 2007).

¹⁰ Use cases describe the interaction between a primary actor (e.g., a warfighter, an S&T manager) and a system being evaluated, with the interaction being represented as a sequence of simple steps.

assessment (chapter 3), an ROI composite index (chapter 4), the probability of achieving an ROI (chapter 5), and a discussion of PPM tools (chapter 6). Chapter 7 will discuss how the use case analysis will be approached in a follow-on study. Finally, the paper will close with a summary of the paper and some concluding remarks (chapter 8).

Chapter 2. Difficulty with Assessing a Return-on-Investment Effect in the Military

Assessing the benefit of Army S&T investments is a difficult task. The issue has been an ongoing challenge for the Army for decades. In addition to difficulties with quantifying the future benefit of an S&T investment, there are many external influences that may shorten the period of technical advantage or add complexity to realizing the envisioned military capability. For example, the globalization of technology may significantly reduce the number of years the U.S. Army has a capability advantage over an adversary. Additionally, efforts addressing netcentric systems developments will make it extremely difficult to attribute the benefit of a single technology to overall warfighting performance within a complex, adaptive system-of-systems environment.

This chapter will attempt to define the problem faced by an Army S&T manager in assessing the benefit of his programs and projects to the Army. It will do this by first briefly discussing the ROI approach within the commercial sector, and then briefly defining the problem faced by the Army. The chapter will then provide a brief introduction of approaches for assessing the benefit of S&T programs. These approaches will be discussed in greater detail in follow-on chapters.

Commercial ROI

Commercial ROI approaches are well known and are critical for many corporate business decisions. Commercial ROI is typically measured in terms of percentage of profit (or loss) for a particular investment, time saved, productivity gained, sales and market share, or customer satisfaction and loyalty. Additionally, ROI analysis can be measured in terms of goals for a specific program. For a safety program, measures of accident frequency rates, accident severity rates, and first-aid treatments would be of interest. For six sigma¹¹ programs, defects, rework, response time, cycle time, and costs are important. And for wellness/fitness programs, a corporate manager might want to track turnover, medical costs, accidents, and absenteeism.

Probably the most widely used and simplest ROI metric is the percentage of money gained or lost on an investment relative to the amount of money invested. In mathematical form, this can be represented as:

$$\text{Simple ROI} = \frac{\text{Gains} - \text{Investment Costs}}{\text{Investment Costs}} \times 100\% = \frac{\$700,000 - \$500,000}{\$500,000} \times 100\% = 40\%$$

This calculation assumes a direct investment relationship to outcome, and that costs are known. It does not deal well with indirect or allocated costs. In addition to this calculation, corporate decisionmakers will also want to address risks and cost drivers, as well as the time needed to

¹¹ Six Sigma is a set of practices originally developed by Motorola to systematically improve processes by eliminating defects.

realize the ROI. As in figure 1, corporate cash flow is very much dependent on reducing costs, including initial investments, as well as accelerating the timing and increasing the magnitude of the ROI.

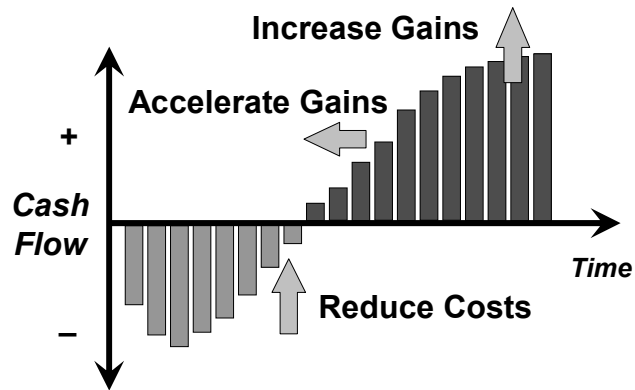


Figure 1. Cash flow analysis with ROI

To keep investors satisfied, corporate decisionmakers are very interested in realizing investment gains (profits, market share, etc.) in the near term—within a year. This near-term attitude has forced many companies to abandon internal strategic technology development expenditures and capital investments. Thus, ROI analysis appears to reduce innovation in a company.¹² However, long-term survival still motivates companies, and one approach to identifying new ideas is to outsource research efforts, including funding universities to accomplish basic research.

The Problem for the Army

Applying this commercial corporate view of an ROI methodology to the military is difficult. The Army doesn't operate with the intent of making a profit, saving money, or expanding its market share. Army S&T is primarily focused on enhancing capability and reducing technical risk. S&T investments may provide benefit in terms of cost avoidance, not to save money, but rather to be able to buy more capability. The Army wants near-term gains to support current operations, but is also concerned about enhancing mid- and long-term capabilities. This latter concern complicates the analysis of the benefit of an S&T effort, in that the benefit may not be seen for many years, and thus may be affected by the globalization of technology, unforeseen changes in the threat, and other uncontrollable outside influences. To add more complexity to the issue, an S&T effort may morph into other technologies or become a critical component of more than one system.

Thus, rather than an ROI approach, it is more appropriate to pursue a cost-benefit analysis. Just as the benefit will be difficult to quantify, the cost is also hard to define. It may be very difficult to capture Army-specific S&T investments and relate them to specific benefits. The S&T investment/resource process is complex. For example, a project may include multiple investment programs/funding sources available to the Army (e.g., Army S&T, teaming with other Services on an S&T effort, teaming with the Defense Advanced Research Projects Agency or other DOD Agencies with S&T efforts, competing for Office of the Secretary of Defense funding, and

¹² 4th Generation R&D; Managing Knowledge, Technology, and Innovation, 175–176.

teaming with agencies outside DOD (e.g., Department of Homeland Security). Additionally, there are multiple reimbursement processes/charging policies across the Army.

It is tough to conduct an ROI assessment in the Army, and it is more difficult because S&T is hard to track in systems. As seen in previous CTNSP studies,¹³ it may be difficult to identify in a system the contributions of Army S&T successes versus industry and academia successes. It is even more difficult to track the contributions of academia, since most research work is not system specific. Collaborations with international universities and industry add even more difficulty to sorting out the contributions to the success of a particular technology. One possible approach to resolving these problems may be to look at total investment in technologies immediately before transition to a given system, but this type of tracking does not normally exist in technology development and system development documentation. This may force one to make an educated guess from time to time.

Recent Efforts

As discussed above, assessing the ROI or benefits of military S&T programs has always been a perplexing problem. There have been many efforts to assess the value of S&T investments, including assessments by subject matter experts and nationally recognized committees, as well as technology reviews, such as the Technology Area Reviews and Assessments sponsored by the Director, Defense Research and Engineering.

An author (Chait) and contributor (Killion) attempted to address the problem in 1994. Their unpublished paper appears in appendix C and identifies three approaches:

1. The first approach was the calculation of a benefit-to-cost ratio (BCR). The BCR is calculated by dividing the rate of return per year by the total investment. This approach is used to determine the rate at which one can expect to recover the initial research investment. As an example, a BCR of 0.5 would mean that the investment is recovered over a 2-year period. Since the Army does not “profit” from the products of basic research (at least in monetary terms), measurement of fiscal benefit were instead made in terms of projected cost savings/avoidance (e.g., in terms of reduced operating costs). This approach is a good example of “cost avoidance” analysis, which will be discussed in the next chapter. However, this BCR approach was very conservative, in that it did not account for non-fiscal benefits such as enhanced military capabilities.
2. The second approach they considered was the “Jimmy Stewart Test.” That is, just as in the movie “It’s a Wonderful Life,” where Jimmy Stewart as George Bailey had the opportunity to witness what would have happened if he had not existed, what capabilities would we not have if the Army’s basic research investment had not been made? In particular, what differences have the products of these investments made in our current

¹³ Richard Chait, John Lyons, and Duncan Long, “Critical Technology Events in the Development of the Abrams Tank: Project Hindsight Revisited,” *Defense & Technology Paper* 22 (Washington, DC: Center for Technology and National Security Policy, December 2005). Also see Richard Chait, John Lyons, and Duncan Long, “Critical Technology Events in the Development of the Apache Helicopter: Project Hindsight Revisited,” *Defense & Technology Paper* 26 (Washington, DC: Center for Technology and National Security Policy, February 2006). Finally see Richard Chait, John Lyons, and Duncan Long, “Critical Technology Events in the Development of the Stinger and Javelin Missile Systems,” *Defense & Technology Paper* 33 (Washington, DC: Center for Technology and National Security Policy, July 2006).

(or future) military capabilities? However, this assessment looks backwards in time and may not be useful for determining an investment strategy.

3. Finally, they consider key scientific measures of merit, which provide an indicator of the impact that Army research investment is making within the broader scientific community. Measures such as numbers of published papers and patents are two examples of key indicators that can be benchmarked via comparisons with industry or other government agencies. In many cases, this approach has great merit. However, the number of papers or patents may not directly relate to the quality of research being performed.

Solving The Problem is Not Impossible

Although the prospects for conducting a precise ROI assessment may look bleak, a cost-benefit analysis is not impossible. The experimentation and acquisition communities have conducted cost-benefit analysis and trade-off analysis efforts for many years, utilizing tools such as simulations to support the analysis. The Army has many unique cost-benefit considerations that may be addressed by the S&T community, though they may be difficult to quantify. For military assessment approaches, benefits may be measured in many terms, including:

- Capability enhancements (e.g., capability advantage over opponents, ability to support current/future doctrine/tactics)
- System effectiveness enhancements (e.g. increasing the enemy:friendly loss exchange ratio)
- System performance enhancements (e.g., accuracy, speed, bandwidth rate)
- Improvements to human performance (e.g., efficiency, which includes time needed to accomplish a task?)
- Cost avoidance
- Productivity improvements
- System quality improvements
- Reductions in risk (performance, schedule, and cost)
- Impact on reliability, maintainability, usability, supportability, affordability, and manufacturability.
- Satisfaction of testers, warfighters, Congress, media, and other identified stakeholders

Discussion of the use and measurement of most of these benefits continues in the next chapter.

Chapter 3. Cost-Benefit Analysis

The first component of the methodology is the cost-benefit analysis. In this assessment, benefits are defined and measured for a given cost. As much as possible, the benefits are measured using data which is known, extrapolated from historical data, or derived from experimentation. When none of this data is available, then estimates from subject matter experts may be a last recourse.

As mentioned in the previous chapter, there are many military benefits that may be useful in a cost-benefit analysis. This paper identifies the most important benefits and groups them into five categories:

- Capability enhancements
- Cost avoidance
- Productivity
- Risk reduction
- Stakeholder confidence

These five categories will be discussed below, followed by a discussion of approaches for evaluating them.

It should be noted that for all of these categories, the assessment of capability enhancements will be difficult when considering the benefits of basic research. By regulation, most basic research does not support a specific military system, group of systems, or military process.¹⁴ A recent National Academy of Sciences study found that in some cases basic research does focus more on needed capabilities, but that it should be “farsighted high-payoff research that provides the bases for technological progress.”¹⁵ If basic research does not have specific applications in mind, it may be very difficult to assess specific enhancements or benefits to future capabilities. As stated earlier, previous CTNSP studies¹⁶ found it difficult to identify the relative contributions of Army and academic basic research to the capabilities of the tank, helicopter, and missiles that were evaluated. As an example of how this can be difficult, research in the area of networking algorithms may be of benefit to mobile, ad-hoc networks, wireless networks, sensor networks, or network integration. In turn, advances in these areas may be of benefit to small-unit operations, Future Combat Systems operations, or even joint operations. The translation of network algorithm research into a measurable military capability would be very difficult. When considering basic research, all assumptions should be documented and considered in detail when measuring its causal effects on military capability enhancements.

¹⁴This is addressed in DOD Financial Management Regulation, DOD 7000.14-R, Vol. 2B, Ch. 5, “Research Development, Test, and Evaluation Appropriations”. The regulation states that “Basic research is systematic study directed toward greater knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific applications towards processes or products in mind,” available at <http://www.defenselink.mil/comptroller/fmr/02b/02b_05.pdf>.

¹⁵ National Research Council Committee, *Assessment of Department of Defense Basic Research* (Washington, DC: National Academies Press, 2005).

¹⁶ See “Critical Technology Events in the Development of the Abrams Tank: Project Hindsight Revisited,” “Critical Technology Events in the Development of the Apache Helicopter: Project Hindsight Revisited,” and “Critical Technology Events in the Development of the Stinger and Javelin Missile Systems.”

Categories of Military Benefits

Capability Enhancements

Given the business of the Army, the first priority goal of most S&T investments is to enhance warfighting capabilities. These capabilities support the conduct of the full spectrum of military operations (offense, defense, stability, and support) and take into account individuals, organizations, and equipment. These capabilities range from as simple as the ability of a soldier to fly a micro-UAV, to as complex as the ability of a Future Combat Systems (FCS) Brigade Combat Team (BCT) to participate in joint urban operations. Given today's focus on the human dimension of netcentric operations, priority should be given to capabilities that relate to human and organizational performance. Finally, the current transformation of the Army makes it imperative to assess capability enhancements to joint operations. An in-depth study of this last assessment area will be presented in a separate CTNSP publication (summarized in appendix D).¹⁷

What has been described is a plethora of warfighting capabilities, the analysis of which would be overwhelming in both time and resources. Thus, to conduct an efficient cost-benefit analysis, the capabilities must be narrowed down in terms of importance and relevance to the S&T projects being evaluated. To accomplish this, the S&T analyst can:

- Meet with warfighters to discuss current capability needs (a capability “pull”)
- Meet with warfighters to discuss future capabilities which may be generated by an S&T effort (a technology “push”)
- Review capability needs identified by the Army, e.g., in published capability development documents or by the Army Capabilities Integration Center (ARCIC). The ARCIC “... designs, develops, integrates, and synchronizes force capabilities for the Army across the DOTMLPF [doctrine, organization, training, materiel, leadership and education, personnel, and facilities] imperatives into a Joint, Interagency, and Multinational operational environment from concept through capability development.”¹⁸

Cost Avoidance

A consistent goal of all S&T investments should be to avoid unnecessary costs and get the best warfighting system/system of systems (SoS) capabilities for the amount of available funding. The Army will not necessarily save money, but rather get more of something (capability, quantity, etc.) for its money. Cost avoidance can be realized in four ways: 1) avoidance of unnecessary costs in the execution of an S&T project, 2) leverage of others' resources, 3) development of the S&T effort with the idea that the cost of its eventual use will be minimized, and 4) development of an S&T effort with the goal of reducing the cost of a fielded system or SoS, a developing system or SoS, or a process. The first two will be the responsibility of the S&T manager. The third and fourth should be considered for cost-benefit analysis efforts. The BCR analysis in appendix C is a good example of a cost avoidance analysis.

¹⁷ Charles Barry, “Army S&T Investment Return in Joint, Coalition, and Civil Agency Interoperability,” forthcoming *Defense & Technology Paper*, Center for Technology and National Security Policy, Washington, DC.

¹⁸ Brigadier General Peter Palmer, U.S. Army, “Army Concepts to Capabilities,” presentation at University of California, Santa Barbara, February 13, 2008.

In assessing cost avoidance, it is important to consider the entire life cycle costs of a system/SoS. This assessment must take into account affordability issues like initial development and manufacturing costs, fielding and training costs, reliability, supportability, and maintainability.

Productivity

The definition of productivity for this analysis does not include the warfighting capabilities discussed above. For example, an S&T effort that enhances the decisionmaking capability of a warfighter is not a productivity improvement, but rather a warfighting capability improvement. Productivity, as considered in this study, relates to S&T efforts that support MANPRINT (Manpower and Personnel Integration)¹⁹ initiatives, enhance military system/SoS development, exploit and enhance the expertise and knowledge of government laboratories, or improve the manufacturability of systems.

The MANPRINT program strongly supports the enhancement of productivity in the Army. Three of its domains are concerned with minimizing the number, skill levels, and training requirements of personnel needed to operate a system/SoS. A fourth domain focuses on human factor engineering concerns to ensure that warfighters and systems are well suited to each other in terms of physical (e.g., space available, required strength of the human, stress on human physiological status) and cognitive (e.g., information displays, information overload) system/SoS design considerations. Usability is included in this domain. Two of the remaining domains, system safety and health hazards, not only support keeping warfighters on the job, but also minimize costs related to payments for job-related physical disabilities and long-term health issues. The final domain, soldier survivability, is more typically identified as a warfighting capability.

Opportunities exist to greatly enhance system/SoS development in today's very high-tech Army. For example, given today's software intensive SoS, such as FCS, great benefits can be realized by S&T efforts that address the interoperability of software so as to make software programmers more efficient in linking together the complex software components of Army weapon and information systems.

The vast expertise and knowledge of the Army laboratories should be exploited by future technology development efforts. Not utilizing this capability disregards a resource that could greatly enhance the military relevance and capabilities of a given technology. Productivity is enhanced because the laboratories need not spend time on learning about combat utility, interoperability, and other military related issues. At the same time, productivity improvements are gained from this interaction as it continues to enhance laboratory expertise and knowledge.

Enhancements in manufacturability will not only improve the productivity of manufacturing facilities, but should also contribute to avoiding some system/SoS life cycle costs.

¹⁹ MANPRINT is the name of the Army's human-system integration program. For more details, see <http://www.manprint.army.mil/programs.html>.

Risk Reduction

Inherent in any system/SoS development effort are risks associated with technical and operational performance, program schedule, and program costs. Technical and operational performance risks are highly related to the technical maturity of the technologies and sub-systems being integrated into a system/SoS. The reduction of technical and operational risks helps to minimize schedule and cost risks by helping system/SoS developments stay on-track with program plans and reducing the need for additional technology development costs to address maturity and capability issues of “problem” technologies. For example, the technical and operational performance of a micro-UAV may be highly dependent on a low-weight, low-volume, low-power, obstacle-avoidance technology. Most likely, this technology is on the critical path of the micro-UAV development plan, so any problems with the technology will immediately affect program schedules and costs. An S&T effort that increases the probability of success of achieving an obstacle-avoidance capability could measure its impact in terms of a percentage of technical, schedule, and cost risks.

Stakeholder Confidence

A final important goal during S&T investment should be to ensure that stakeholders are satisfied with their investment selections and their development progress. Many stakeholders are involved with S&T efforts, but the number and interest of stakeholders normally increases with budget activity (BA) levels. For basic research activities, stakeholders of interest may primarily be the project managers of applied research projects and the laboratory director. For the most mature S&T efforts—advanced technology development efforts—stakeholders may include the laboratory director, program managers of systems targeted for technology insertion, warfighters (especially those involved with developing advanced concepts and requirements), and perhaps even congress and the media.

Assessment of Categories of Military Benefits

There is no standard assessment approach for the evaluating utility in respect to categories of military benefits. This is the most complicated type of ROI assessment. However, given today’s modeling and simulation tools, this type of assessment is possible and even quantifiable. The assessment of cost avoidance, productivity, risk reduction, and stakeholder confidence are relatively easier, with histories of proven assessment already evident within the military.

Capability Enhancements

The assessment of military capability enhancements can best be accomplished with the use of experimentation. An experiment can demonstrate basic system capabilities and confirm the system’s value in accomplishing given objectives. If the experiment is designed correctly and metrics are developed to track the “cause and effect” of experiment outcomes, the value of an S&T effort can be extracted from the experiment results. Experiments will also provide the freedom to pursue excursions to gather additional data in a particular technology or area of interest.

The first step in designing an experiment is to identify a military utility that can be impacted by S&T efforts. For example, in utilizing a micro-UAV to search a room during urban operations, a soldier operator needs to focus his attention on his surrounding area as well as the actionable information coming from the micro-UAV, rather than concentrating on the mechanics of flying

the micro-UAV. With this military utility in mind, the next step is to identify supporting objectives—e.g., determine the technical and operational benefits of obstacle avoidance technologies in micro-UAVs. Once this is accomplished, sub-objectives are identified for each objective. For example, sub-objectives in the above case might include: identify the impact on flying multiple micro-UAVs in close proximity to one another; identify the impact on the operator’s ability to fly the micro-UAV; etc. The next step is to identify hypotheses under each sub-objective that must be proven true or false—e.g., “at least three micro-UAVs can be flown in the same room;” or “the operator will have a significant (at least 100%) improvement in flying the micro-UAV.” Finally, data is needed to prove or disprove each hypothesis, and that data must be gathered using appropriate metrics. A data collection plan is then developed that is used as a primary driver for designing the experiment, including its scenario and vignettes.

Developing the metrics for an experiment is not a trivial task. Metrics are divided into three levels:

- Level 1: Technical Performance Metrics. Examples of these metrics are detection range, probability of detection, probability of false alarm, and time needed to process sensor information. These metrics can be called measures of performance (MOPs).
- Level 2: System Effectiveness Metrics. System effectiveness metrics may assess the impact of the sensed information on mean acquisition time of obstacles by the micro-UAV, the reaction time of the micro-UAV to miss obstacles, the ability of the military user to “fly” the UAV, and other similar measures. These metrics can be called measures of effectiveness (MOEs).
- Level 3: Operational Utility Metrics. Operational utility metrics assess the value of the technology on the ability of the user to sense targets, the staying power of the micro-UAV, the survivability of the micro-UAV, and other operational measures. These metrics can be called measures of value (MOVs).

The development and documentation of these MOPs/MOEs/MOVs is described in more detail in appendix E.

The design of the experiment should take into account the military utility, objective, sub-objectives, hypotheses, and data collection needs; the resources available; the maturity of the technologies and systems; and the tools available. Much work has been done in the experimentation community, as well as within the test and evaluation community, in the “design of experiment (DOE)” area. The DOE is a structured, organized method for designing experiments to determine how well a technology X performs when assigned tasks Y_1 to Y_n under a set of conditions Z_1 to Z_n .

When technologies are at the conceptual level and not well defined, discovery experiments can be conducted using low fidelity simulations, games, or even thought experiments (a.k.a. Gedanken Experiments). An agent based simulation (e.g., MANA, Pythagorus)²⁰ is an example

²⁰ These two agent-based simulations were used with great success with the Marine Corps program called Project Albert.

of a low fidelity simulation. The advantage of using these simulations is that they are easily and quickly set up and can be run significantly faster than real time, allowing for numerous runs and the use of data farming techniques to identify emerging trends and issues²¹. Modified commercial games with warfighter players may also suffice for these experiments. A thought experiment typically consists of a group of subject matter experts sitting at a table and employing imaginary situations to understand how a technology may contribute to a military capability now, or how it may contribute to future capabilities.

As a technology effort matures and the technology is better defined, higher fidelity constructive (computer-based) simulations may be employed. The Army has a number of computer-based simulations, including:

- At the Brigade and below level
 - Combined Arms and Support Task Force Evaluation Model (CASTFOREM)
 - JANUS
 - One Semi-Automated Forces (OneSAF)
 - Combat XXI
- At the Corps and Division level
 - Vector-In-Command (VIC)
 - Eagle
 - Advanced Warfighting Simulation (AWARS)
- At the joint warfighting level
 - Tactical Warfare (TACWAR)
 - Joint Warfare System (JWARS)

In these constructive simulations, computer-generated people operate computer-generated systems. Real people stimulate (make inputs) to such simulations, but are not involved in determining the outcomes. More information on these simulations, their capabilities, and other Army simulations are normally available on the Web.²² The U.S. Army Training and Doctrine Command (TRADOC) Analysis Center (TRAC) routinely uses these simulations for conducting experiments for assessing DOTMLPF issues.

As technology development continues further, and mock-up systems can be developed, the use of virtual simulations is possible. A virtual simulation involves real people operating simulated (or mock-up) systems. Virtual simulations inject human-in-the-loop in a central role by exercising motor-control skills (e.g., flying an airplane), decision skills (e.g., committing fire control resources to action), or communication skills (e.g., as in communicating with other members of a team).

When prototype technologies and systems are available, live simulations can be used. In live simulations, live people operate real (or near-real) systems. The Army Battle Labs support a significant amount of live experimentation for the Army. They also have constructive and virtual simulation capabilities.

²¹ Data farming is much different than data mining. The latter merely searches for data that already exists. The former stimulates (or seeds) a simulation for creating data that can be harvested for analysis.

²² An example of the available information can be found at, see <http://www.amso.army.mil/resources/smart/pol-guid/guidance/ref-guide/sec-VI/tools.htm>.

Finally, when conducting an experiment that has many technologies and systems of various maturity levels, and when resources (people, equipment, ranges, funding) are limited, it may be necessary to conduct integrated live-virtual-constructive simulations. In doing so, it is also possible to conduct an integrated live-virtual-constructive experiment in a distributed manner—in which live-virtual-constructive participants are geographically separated by a few or even thousands of miles. The Army Battle Labs are linked together to enable the execution of distributed experiments.

In the past, the Army S&T community had to integrate their technologies in live Army Warfighting Experiments conducted at the Battle Labs. With the availability of today's games and simulations, as well as networking capabilities, it is possible for the Army S&T community to conduct experiments in-house, or as a participant in a distributed simulation. To support this capability, the Army initiated the modeling architecture for technology, research, and experimentation (MATREX) program, which is the foremost distributed modeling and simulation environment in the US Army. It provides a unifying M&S architecture, supporting tools, and infrastructure to ease the integration and use of multi-resolution live, virtual, and constructive (LVC) applications.

In addition to collecting data in these experiments for individual MOPs, MOEs, and MOVs; the following measures could be considered for an overall assessment of a particular S&T program:

- Number of capabilities that are achievable with the S&T program
- Number of capability needs (across the full spectrum) identified by warfighter (e.g., the ARCIC) that are being addressed by the S&T program
- S&T program's contribution to the ability of Army to operate in a joint environment (for additional information, see separate Defense and Technology Paper²³ on this subject)

Cost Avoidance

As stated earlier, cost avoidance assessments can be conducted in a manner similar to that performed by Chait and Killion in their unpublished paper in appendix C.

In addition to the process described in appendix C, the following issues should also be addressed (if possible):

- Additional costs that might occur without the S&T (this is the “predictive” version of the “Jimmy Stewart Test”)
- Costs that may be avoided because of improved efficiencies
- Technology and system investment decisions impacted by an S&T effort
- Impact of an S&T effort on the Army Strategic Plan, which includes
 - Investments in systems, programs, etc.
 - Leveraging of investments by others

Productivity

MANPRINT assessments have been common in the Army since the early 1990s. They are based on a number of assessment techniques, including the use of simulation tools, like the Improved Personnel Research Integration Tool (IMPRINT) and JACK. The assessment processes are well

²³ See “Army S&T Investment Return in Joint, Coalition, and Civil Agency Interoperability.”

defined and found in Army documents like the *Manpower and Personnel Integration MANPRINT Handbook*²⁴ as well as the following documents found on the MANPRINT web site²⁵:

- Defense Acquisition Guidebook v4.5, with multiple views and download options for the guidebook
- Army Regulation 602-2: Manpower and Personnel Integration (MANPRINT) in the System Acquisition Process.
- Army Regulation 602-1: Soldier-Materiel Systems, Human Factors Engineering Program
- Army Regulation 385-10: Army Safety Program
- Army Regulation 385-16: System Safety Engineering & Management

The assessment of productivity improvements to system/SoS development and manufacturing should be conducted closely with the cost avoidance assessments. Many of these productivity improvements will overlap in content, but have unique benefits like shortened times, enhanced products, etc.

Risk Reduction

Risk reduction (a.k.a. risk management) efforts are often addressed by program managers in their development plans, and are rigorously addressed in test and evaluation activities. The Defense Acquisition University has published a Risk Management Guide that provides guidance and tools for risk management, such as risk rating matrices for conducting risk assessments. Such guidance may not be applicable to S&T efforts, so qualitative assessments by subject matter experts, experienced program managers, and customers may be necessary.

Stakeholder Confidence

Normally, this assessment is conducted using a combination of surveys, interviews, and presentations to stakeholders. Interaction with integrated product teams may also be useful. It should be pointed out that the recent use of contractors as lead systems integrators may add complexity to this assessment, because of proprietary concerns and the transfer of intellectual property. Congress recognizes the complexity of using contractors as lead systems integrators and has noted concern.²⁶

As part of the surveys and interviews, data should be collected to attempt to answer the following measures (note that these are based on historical information, and not on anticipated outcomes of current S&T efforts):

- Number of programs retained/terminated/modified because of S&T investments
- Number of program management offices whose focus on mission capabilities was clearer because S&T successes

²⁴ Manpower and Personnel Integration (MANPRINT) Handbook, available at <http://www.manprint.army.mil/manprint/docs/MPTHandbook.pdf>.

²⁵ MANPRINT website, available at <http://www.manprint.army.mil/home/references/manprint.html>.

²⁶ William Mathews, "An End to Lead Systems Integrators," *Defense News*, 10 December 2007.

Completing the Cost-Benefit Analysis

Once the above assessments are completed, the results can be tied to the cost of a particular S&T effort, thus allowing the S&T manager to rank his projects in terms of degree of benefit per amount of cost. Comparisons of both similar and dissimilar projects can be used to conduct trade-off analyses that support the end goal of getting the most military benefit for a given cost.

The cost-benefit analysis just discussed may generate a long list of calculated benefits, which may make pair-wise comparisons very complex. The next chapter recommends an approach to grouping benefits for a simpler comparison.

Chapter 4. Composite Index

If a thorough ROI-like analysis is conducted, the S&T manager will probably be faced with a long list of calculated benefits, which may present serious difficulty in making pair-wise comparisons among various S&T projects. One way to make the comparison less complex is to develop a composite index that groups similar benefits together, and also weights each benefit based on its importance to the Army. For example, the benefits composite index for stealth technologies ($I_{B\text{-Stealth}}$) can be represented as:

$$I_{B\text{-Stealth}} = W_1(P_1/R_1) + W_2(P_2/R_2) + W_3(P_3/R_3) + W_4(P_4/R_4) + \dots + W_n(P_n/R_n) \quad [1]$$

or

$$I_{B\text{-Stealth}} = W_1(BCR_1) + W_2(BCR_2) + W_3(BCR_3) + W_4(BCR_4) + \dots + W_n(BCR_n) \quad [2]$$

Equation 1 would be used when requirements for a technology are known. For stealth technologies, these requirements may be described in terms of acoustic, visual, infrared, and electro-magnetic signatures, with the goal as the smallest signature possible. Given this need to reduce signatures, the overall goal would be to achieve the smallest index possible. The requirements will be listed as ($R_1 \dots R_n$), with, as an example, R_{acoustic} equal to 10 decibels.²⁷ The technical performance of each S&T effort is represented as ($P_1 \dots P_n$), with P_{acoustic} equal to whatever is measured in a laboratory. In this stealth example, the closer a P_1/R_1 value is to zero, the better it is. Thus, given a requirement of 10 decibels, a value of 6 decibels in performance would be better than a value of 8. It would not be correct to include an R_x in this example, as that would imply that the higher the value, the better the technology. It would be correct to include such a requirement in an index whose value is being maximized.

For both equation 1 and equation 2, the weights of each requirement ($W_1 \dots W_n$) are developed with customer or user input. The input should be derived from personal experience or some form of experimentation. For example, for a micro-UAV, the user may put a high weighting on noise signature, a moderate weighting on visual signature, and a low weighting on infrared and electro-magnetic signatures. The weight values must be determined with respect to the importance to the customer or user. In some cases simple values of 1 (low), 2 (moderate), and 3 (high) may be sufficient. For technologies that have a significant impact on customer or user needs—for example, countering improvised explosive devices—the weighting may need more rigorous values of 1 (low) through 10 (high).

When requirements are not known, but benefits can be calculated, as discussed in Chapter 3, and there are many benefits for pair-wise comparison, it may be advantageous to use equation 2. In this equation, the BCR_x term is used instead of the P_x/R_x term. The BCR_x term represents a benefit identified in Chapter 3. Each BCR_x term may be a capability, cost avoidance, risk reduction, customer satisfaction, or any other benefit identified by the S&T manager. It must be quantified so that it makes sense in the equation; if possible, normalize all the BCR_x terms. In

²⁷ As a comparison, a whisper in a library is 30 decibels and a telephone dial tone is 80 decibels.

this particular equation, the program manager will play just as important a role as customers and users in developing the values of the weighting factors.

Each index should be defined in detail similar to measures of performance and measures of effectiveness (see appendix E), with documentation as described in the example in the box below.

Title of the Index. Weighted Stealth Index ($I_{B-Stealth}$) Definition of the Index. $I_{B-Stealth}$ is a weighted sum of the performance of various stealth technologies in relation to defined requirements. Input data are the relative weights of each requirement ($W_1 \dots W_n$) and the performance ($P_1 \dots P_n$) observed in each requirement ($R_1 \dots R_n$). $I_{B-Stealth} = W_1(P_1/R_1) + W_2(P_2/R_2) + \dots + W_n(P_n/R_n) = \sum \{W_i(P_i/R_i)\} \text{ for } i = 1 \text{ to } n$ Examples of performance requirements are: acoustic, visual, infrared, and electromagnetic signatures, each of which is measured directly or rated by evaluators on a common scale. Notation of the Index. A weighted sum expressed as a real number. Limitation/Range of the Measure of the Index. The magnitude of the $I_{B-Stealth}$ depends on the values defined for W, P, and R. The maximum value is found by assigning for each “n” term its appropriate $W_{max} (P_{max} / R_{max})$. Reasoning for the Index. The $I_{B-Stealth}$ is intended to combine performance of all relevant requirements to preclude over-valuing of some requirements in pair-wise comparisons. Relevance of the Index. The measure can be used to compare various stealth technologies. Variations of Use of the Index. The relationship of an individual stealth technology to other stealth technologies may be evaluated by attempting to control the effects all other stealth technologies, or by eliminating some stealth technologies. Associated Indices or Metrics: Size-weight-power index; measure of performance of micro-UAVs in avoiding collision with other micro-UAVs; measure of effectiveness of operator. Linkage to Metrics in Army S&T Programs: Identify linkages to metrics in Army Technology Objectives (ATOs), Army Technology Demonstrations (ATDs), and Technology Transition Agreements (TTAs)
--

These last two chapters have identified approaches for assessing the benefits of a given S&T investment. However, the approaches so far do not address the likelihood of achieving the stated benefits. This problem will be addressed in the next chapter.

Chapter 5. Probability of Achieving the Anticipated Benefit

Much thought and energy goes into the identification of S&T efforts to be pursued in Army laboratories and engineering centers. Not surprisingly, once the ROI or benefit analysis is completed, and the results are provided either individually or as a composite index, numerous technology efforts will often receive ratings that are not only high, but similar to each other. Since there are inherent uncertainties and errors in the assessment methods described in the previous two chapters, it would be beneficial to further scrutinize S&T efforts, to ensure that the right projects are selected for investment.

To accomplish this additional assessment, a method is proposed for calculating the probability of achieving the ROI or benefit. For example, if based on its current schedule a technology will fully mature in 2012, but is needed for insertion into a military system in 2010, the receiving system's Program Manager (PM) may be forced to use something in lieu of the technology until it becomes available. Otherwise, the PM may have to drop it from his system integration plan altogether. If the latter happens, the probability of transition will be near zero. Therefore, the probability of achieving the anticipated ROI (P_{ROI}) or BCR (P_{BCR}) will be near zero—assuming no other systems can use this particular technology. So investing in this project would be fruitless, unless its schedule for completion can be changed.

The development of an overall P_{ROI} or P_{BCR} should consider many factors. In this paper, technical success, transition opportunity, technology insertion, funding, and ability to maintain an advantage will be considered. The overall probability will be the product of the probabilities for these factors. An S&T manager may identify other factors, and can incorporate them in the probability analysis using the methods described in this chapter.

Unfortunately, as will be discussed below, the probabilities derived for most of these factors will be based on qualitative judgment, especially when considering basic research efforts. Considering this, the initial reaction might be to disregard this assessment. However, the thought process required to derive the probabilities will increase the S&T manager's confidence level in selecting the most appropriate projects for a given amount of investment resources.

Components of the Probability

Technical Success

The first factor that must be considered is the project's probability of technical success. This factor is usually one of the top selection criteria in evaluating responses to broad agency announcements or requests for proposals. The source selection is usually accomplished by government personnel who have subject matter expertise in the technology under consideration. As an aside, this is one of the strongest arguments for maintaining the technical expertise of government personnel—a.k.a. maintaining smart buyers. The assessment of technical success is usually based on the technology readiness level (TRL) of the initial concept/technology, the anticipated TRL goal, and the ability of the performer.

S&T managers have the in-house expertise to conduct such an assessment and should do so for each S&T effort being considered for investment. From this assessment, a probability of

technical success should be derived. This probability should be made with regard to time and availability of funding. For example, when assessing the probability of developing autonomy in robotic systems, there may be a 50% probability of achieving complete autonomy for micro unmanned air vehicles in 2015 and a 60% probability of achieving complete autonomy in 2025. However, an increase in the level of funding may change the probabilities to 75% and 90%, respectively. A plot of this information, similar to figure 2, would be very useful in making comparisons of investment strategies.

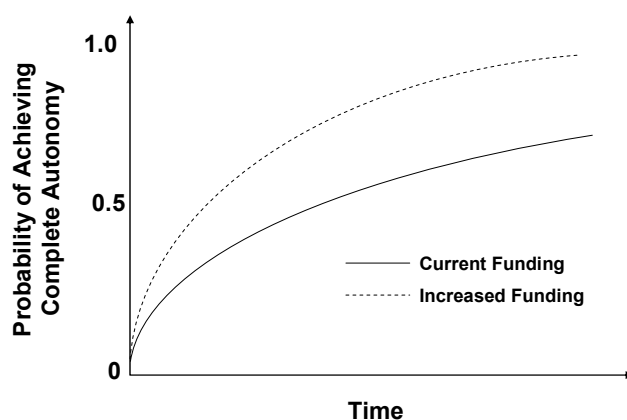


Figure 2. Autonomy versus time and funding

Transition Opportunity

The movement of an S&T effort from one research and development BA level to the next higher level activity is called “transition.” The budget activities²⁸ of interest to S&T managers for transition are:

- BA 1, Basic Research
- BA 2, Applied Research
- BA 3, Advanced Technology Development
- BA 4, Advanced Component Development and Prototypes
- BA 5, System Development and Demonstration²⁹

An S&T effort can be transitioned if the following conditions are satisfied: there is an identified customer, the effort is being developed to satisfy the needs of the customer, it is mature enough at the time of transition for the customer to take over further development, the customer has adequate funding in place to accept the technology, and adequate S&T knowledge is provided to the customer. This last factor can be accomplished by keeping the customer’s technology experts involved with the S&T effort prior to transition, by transitioning adequate knowledge (e.g., documents), or by transitioning subject matter experts with the S&T effort to work with the customer until he is able to take on complete development responsibility.

²⁸ Previously, DOD budget categories were listed as 6.1 (basic research), 6.2 (applied research), etc.

²⁹ A description of each of these budget activities can be found in DOD Financial Management Regulation, DOD 7000.14-R, Vol. 2B, Ch. 5, “Research, Development, Test and Evaluation Appropriations,” 5-2–5-3. The full text is available at <http://www.dod.mil/comptroller/fmr/02b/>.

A transition plan for the S&T effort should include all of the above options. Based on the status of each of the items, the S&T manager should be able to make an assessment of the probability of transitioning a specific S&T effort.

Technology Insertion

The end goal of an S&T development effort is to insert a technology as a component of a future fielded system or capability. Based on the S&T transition plan mentioned above, and the ability of the BA 4 and BA 5 customer to fully develop the technology and its targeted system, a time for technology insertion may be estimated. The timeliness for technology insertion is critical, for if it is too late, it will not be part of the production line for the fielded system, and may or may not be part of future system upgrades, depending on available funding and system design.

For example, assume that an FCS BCT is expected to be designed and initially fielded in 2014, with Spin Out 1 (Unattended Munitions) in 2008, Spin Out 2 (UAVs) in 2010, and Spin Out 3 (Unmanned Ground Vehicles) in 2012.³⁰ Therefore, the FCS program office may want an obstacle avoidance technology transitioned to BA 4 by 2008, in order for it to be fully transitioned to BA5 in 2009, and included in Spin Out 2 in 2010. If the obstacle avoidance technology is ready by 2008, it will have a 100% probability of being inserted into the FCS BCT. However, if it is not ready until 2009, it may have a 50% probability of insertion; and if available in 2010, it may have a 10% probability of being part of the FBCT in 2014.

Besides time, other factors may affect technology insertion. Factors such as jointness, interoperability, life cycle costs, and MANPRINT³¹ concerns need to be considered in the final calculation. Finally, an S&T effort does not normally become a fielded system or capability without the integration of other S&T efforts. One final consideration for calculating the probability of technology insertion is the assessment of the synchronization of the effort with other S&T and BA 4 and BA 5 activities.

Adequate Funding

Technical achievements, transition opportunities, and times for technology insertion are all highly dependent on availability of funding. Without funding, work cannot be done. With inadequate funding, schedules will slip. With increased funding, schedules may accelerate. The probability of having adequate funding must be developed given the anticipated technical success, transition plan, and technology insertion point defined above.

Maintaining a Capability Advantage

As mentioned earlier in this paper and in appendix B, globalization has significant effects on technology development—both good and bad. The good aspects of globalization—the global collaboration environment—may positively affect the probability of technical success described above, and should be incorporated in the assessment. On the other hand, globalization may be detrimental in maintaining a capability advantage in that it quickly and broadly disseminates technical information, thus significantly reducing the amount of time that a country may have a

³⁰ Brigadier General Peter Palmer, “Army Concepts to Capabilities”.

³¹ MANPRINT is concerned with factors such as how many personnel will be needed to use the technology, how skilled will the users need to be, what human factor engineering concerns are there, and others. For more details, see <<http://www.manprint.army.mil/programs.html>>.

capability edge. The calculation of this probability should be plotted with respect to time, and then used to extract a specific probability for a particular year.

Calculating the Probability

For a particular S&T effort, the final P_{ROI} or P_{BCR} will be calculated using:

- The probability of technical success (P_{TS})
- The probability the S&T effort will be transitioned (P_{Tr})
- The probability the S&T effort will inserted into a future fielded system or capability (P_{TI})
- The probability that adequate funding will be available to pursue the S&T effort (P_{AF})
- The probability of maintaining a technological advantage by a particular year ($P_{TA/YR}$)

So P_{ROI} or P_{BCR} is defined as:

$$P_{ROI} = P_{TS} \times P_{Tr} \times P_{TI} \times P_{AF} \times P_{TA/YR}$$

If a probability component does not apply, then it should be set to “1.0” and the final probability number should be flagged, noting that the component did not apply. For example, when evaluating a basic research effort, the P_{TI} would not apply, since basic research is not system specific. Additionally, it may be difficult to determine a $P_{TA/YR}$.

Once the likelihood of achieving benefits is addressed, it may be advantageous to view all the benefits, indices, and probabilities together in a spreadsheet or similar tool. This view of the depth and breadth of an S&T program will allow the S&T manager better visualize the pros and cons of each S&T project. However, what is still needed is a tool that compiles all of this information and allows the S&T manager to conduct a trade-off analysis to best determine which projects should be funded, given limited S&T investment resources. An approach for addressing this task is the use of PPM tools, which will be discussed in the next chapter.

Chapter 6. Project Portfolio Management Tools

Another increasingly popular method of evaluating investment options and likely returns is the use of PPM software. PPM software originated in the business world, influenced by the growth of project management as a specialized knowledge area and the increasing use of complex, structured processes to perform it. As the use of business software expanded into project management in the last several decades, users began to look for software to help them not only plan and evaluate projects, but also to select which projects to undertake. As the implementation of a project usually requires some expenditure of resources, or *investment*, PPM tools are created to help decisionmakers estimate the absolute and relative rewards of a series of possible projects. In other words, PPM software is designed to evaluate return on investment.

Ideally, PPM tools should take into account the following four functions:

- *Analysis*: link objectives to vision, goals, priorities, and capabilities; develop performance measures; and identify gaps and risks
- *Selection*: identify and select the best mix of investments to achieve capability goals and objectives across a portfolio
- *Control*: ensure investments within portfolios are managed and monitored to determine whether to continue, modify, or terminate them
- *Evaluation*: measure actual contributions of a portfolio towards improved capability and support adjustments to the investment mix

Many commercial PPM tools focus on the *control* and *evaluation* of the portfolio. Superior PPM software not only assists users with the *Selection* phase of project management, but also facilitates the initial *Analysis* phase, the requirements of the *Control* phase during implementation, and the *Evaluation* phase upon project completion.

Evidence of the value of PPM tools in making S&T decisions is demonstrated by the Office of the DASA(R&T)'s choice to have used an S&T portfolio analysis tool developed by Navigant Consulting in collaboration with the United States Army, since 1999. In addition, the U.S. Army Engineer Research and Development Center (ERDC) uses the tool for its S&T portfolio. Other Army S&T organizations use the process or are in discussions to use it to evaluate their portfolios. The process, known as Portfolio Discovery®, supports a consensus-based S&T evaluation and portfolio analysis that has included representatives from Army laboratories, TRADOC, Program Executive Offices, Program Management Offices, and the National Academy of Sciences' Board on Army S&T. Portfolio Discovery® can be applied to any form of S&T investment, including ATOs, non-ATO-designated S&T, and Small Business Innovation Research (SBIR) efforts. The process has been applied to all stages of investment including 6.1 (Basic), 6.2 (Applied) and 6.3 (Advanced Technology Development), although DASA(R&T) only applies it to 6.2 and 6.3 ATO investments. Portfolio Discovery® is based on a proprietary methodology and consists of evaluative and analytical sub-processes, a proprietary software tool, as well as cultural and behavioral techniques. The primary purpose is to assess the alignment of S&T investments with organizational goals and objectives. The resulting analysis sheds light on

portfolio strengths and weaknesses, alignment with strategic objectives, and describes trends over time. Resource allocation decisions can also be informed by the analysis.³²

The Navigant tool is used within the Office of the DASA(R&T) to ensure ATOs support TRADOC needs, initially focusing on TRADOC's Future Operational Capabilities (about 11) in 1999 and then, in later years, Capability Gaps (about 66). Recently, TRADOC has shifted to Warfighter Outcomes (about 90). The tool has been very successful in past years in aligning ATOs with Future Operational Capabilities and Capability Gaps. However, since the metrics in the tool are directly related to the TRADOC taxonomy of capabilities, this last change to more numerous Warfighter Outcomes may complicate the use of the tool.³³

The Army's apparent satisfaction with Navigant Consulting's current PPM tool demonstrates that institutions managing large portfolios of S&T projects perceive an advantage to using PPM software. In this case, the advantage seems to be in the visual display of the results of Navigant's evaluations of individual project performance, in relation to the agency's strategic goals. Additionally, the individual results are agglomerated into categories of investment, which allows aggregated performance to be compared to cost across various categories of S&T R&D. However, Navigant's tool does not appear to demand or apply a consistent set of criteria across multiple institutions; instead each client is allowed to set its own evaluation criteria. Furthermore, the relationship between the performance scores and any quantitative data on investment performance is not specified or required. The tool appears to allow Navigant and the client to create performance scores without a predefined methodology, or agreement on a formula for incorporating empirical data. A primary goal of this methodology is to create a rigid and consistent framework for S&T evaluation that relies on empirical performance data in predictable relationships to the ultimate scores. Thus, the Navigant tool would require substantial modification in order to enforce our suggested methodology.

We have performed a preliminary evaluation of several other portfolio analysis tools to determine how easily they might be adapted to our analytical methodology. The common goal of these tools is to help decision makers understand the consequences of many possible allocations of institutional investment. "Investment" here refers to money spent in improving the capabilities or performance of the given entity—more similar to a business investment than to a purely financial investment, in that the investment is measured in terms of benefits, rather than monetary return. Each software tool uses a different method to evaluate the benefits and (in some cases) costs of the possible allocations.

The four additional PPM tools discussed in this paper:

- PALMA: a software tool developed by MITRE that provides a graphical and analytic framework to help decision makers select the best combination of investments based on cost and operational benefit. PALMA develops the "efficient" frontier, identifying portfolios (and the elements in each) that provide the most benefit at a specific budget or funding level. It also graphically depicts the capability, mission, and goal to be achieved within a hierarchical decomposition.

³² Email correspondence with John Walker, Navigant Technologies, Inc., June 27 2008.

³³ Conversations with Jagadeesh Pamulapati, Office of DASA(R&T) and John Walker, Navigant Technologies, Inc., June 18, 2008.

- **Balanced Scorecard:** a methodology based on approach developed in the early 1990s by Drs. Robert Kaplan (Harvard Business School) and David Norton. It is not a current tool. It is based on decomposing an organization's vision for the future (e.g., mission, strategic objectives) into focus areas, then critical success factors (tactical objectives) for each focus area, and finally metrics that are measurable, outcome focused, and linked directly to tactical objectives.
- **Equity:** Uses a “balance-beam” method to quantify the value of investments, through facilitated sessions. It allocates resources among alternative investments given budget.
- **Analytical Hierarchical Process (AHP):** a process that uses pair-wise comparison. The AHP was invented by Dr. Thomas Saaty in the early 1970s and developed by Expert Choice, Inc. and The Analytical Sciences Corporation.

PALMA

The PALMA tool prompts the decisionmaker to create a hierarchical map representing the tasks to be performed by a given entity—be that a weapons system, technology area, or human organization. The map is displayed as a pyramidal series of linked boxes, beginning with the primary goal of the entity and branching downward to each objective required to achieve the goal, and further downward to the tasks performed to achieve each objective (see figure 3). At the lowest level of the “tree,” the user assesses each task and inputs a performance score to the corresponding node. After the relative weight of each branch is determined by the user, the performance of any given objective at the next level of the tree is derived from the weighted average of the objective's “branches,” one level below. The performance of each box is color-coded; the default setup links the performance of each node to a traffic light color schema (green signals good performance, red is unsatisfactory) enabling the user to develop a general impression of performance across the system at a glance. When the baseline performance tree has been created, the user goes on to input possible investment options. Every option is described both by cost and by the improved performance it will bring to one or more tasks. Improving the performance of a task on the bottom of the tree will change, in proportion, the average performance score of every connected higher branch. In this manner, the decisionmaker can obtain a visual picture of how different investment options will impact the performance of the object. An alternative choice is to have the software map an “efficient frontier,” a graph that plots the total performance and cost of every possible combination of investment options.

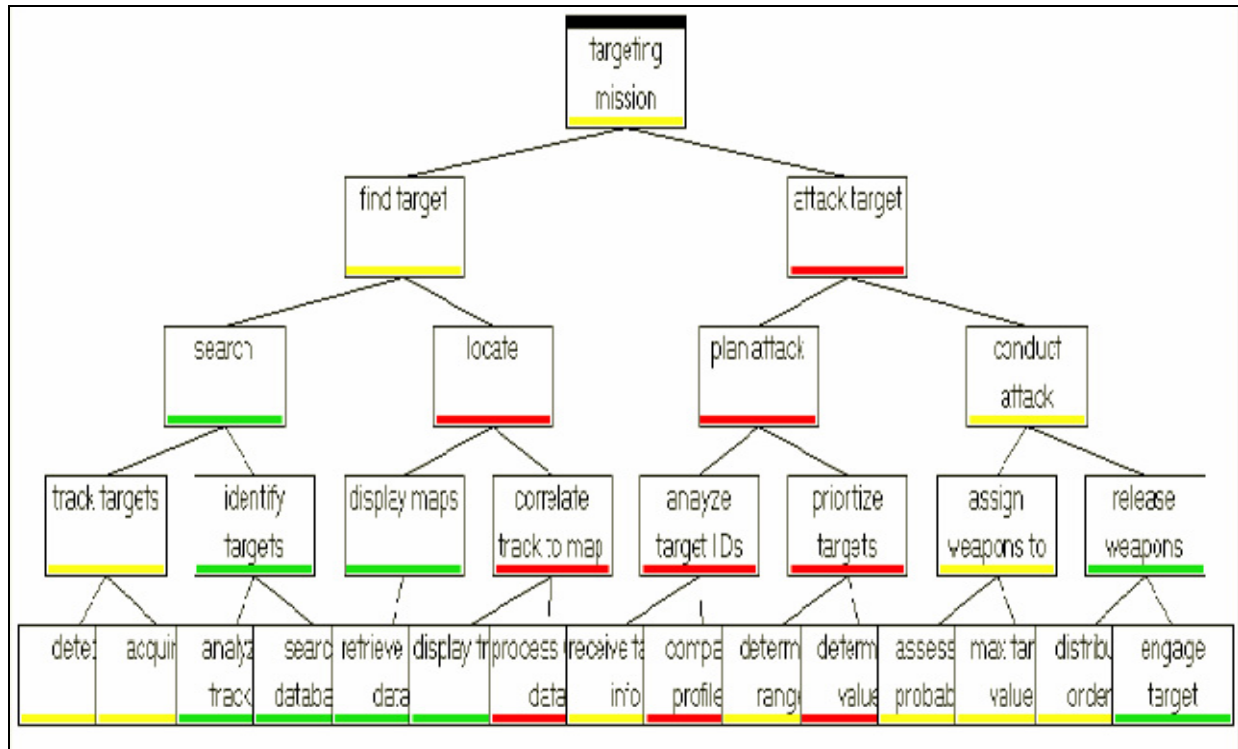


Figure 3. PALMA Strategy-to-Task “Tree” Hierarchy with Color Scheme³⁴

Analysis. The PALMA evaluation is structured in a fashion similar to our methodology, and may thus be relatively less complicated to incorporate than other tools reviewed here. MOPs would be used to fill in the bottom layer of the tree hierarchy. The MOEs and MOVs would represent higher levels of the tree, whose scores should be mathematically related to the MOP results. In some scenarios, the tree potentially could continue upward beyond a single technology area or system component and be used to calculate improved effectiveness gains across multiple systems. The ROI “Composite Index” described in our methodology is very similar to the manner by which PALMA uses weighted averages of a lower-level node set to calculate the next level of the hierarchy. Note that our case study in this report will focus on analysis of MOPs and MOEs, while MOVs will not be considered for this study.

PALMA appears to hold the following additional advantages over other tools evaluated: the clearest visual picture of performance dependencies

- tree hierarchy naturally accommodates MOP/MOE/MOV distinctions
- best customization options (force include/exclude, conditional funding)
- evaluation structure facilitates integrity controls (for example, reviewers might use only empirical data for bottom of the tree, and agree on relative weights before data input)
- tree creation forces thinking about “*Analysis*,” multi-year setup allows “*Control*” and “*Evaluate*”

³⁴ Picture from “*Investment Analysis Using the Portfolio Analysis Machine (PALMA) Tool*.” MITRE Corporation.

On the other hand, some areas of concern are:

- Public documentation does not demonstrate or discuss the mapping of “many-to-many” node relationships. This capability is important to realistically model situations where improving one key metric (or, in our system, MOP) may affect the performance of many entity tasks, or the tasks of many entities. The developers have provided assurances that the software supports this requirement: the effectiveness of the implementation is not known.
- There is no built-in capacity to model negative modifiers to performance, such as might result from a technological “tradeoff” that increased Node X at cost to Node Y. An example would be a developed “Active Protection” armor package that when substituted for passive armor decreased average vehicle damage—Node X—but also increased the likelihood of catastrophic vehicle damage during system malfunction—Node Y.
- It is not clear if ‘rules’ can be set that change the weight of a given node if its performance passes a certain threshold. For example, a 10% decrease in the visual signature of a UAV may only increase effectiveness by 3%, because the vehicle is still observable to most radar instruments, but a 60% decrease in visual signature might cross the ‘effective invisibility’ threshold and increase effectiveness by 200%. Ideally, this kind of variation would be pre-configurable in the software “rollup rules,” but it appears to require manual revision by the user at this time.
- There is no built-in capacity to link the performance of nodes across multiple tree hierarchies, as would be helpful when a single technological investment might “break out” and affect the performance of several systems. MITRE has in the past made available a secondary, external tool for combining the sum of changes in a series of tree hierarchies: the tool’s utility is unknown.
- There is no obvious capability for including the “probability of achieving the ROI.” Adaptations can be imagined, such as the insertion into the hierarchy and scoring of one or more nodes representing ‘risk’, but they may be less clear to the user, or pose implementation difficulties.

Other Evaluated Tools

ExpertChoice® is a project management software platform that relies on the “Analytical Hierarchy Process” to assist project managers in prioritizing the relative advantage of competing alternatives, whether those alternatives are investment options, business strategies, or product choices. The methodology is mathematically complex. In summary, the user is asked to rank the significance of their multiple desired objectives or capabilities on a 1-10 ordinal scale. The significance of each capability is calculated relative to that of every other capability, in a series of “pair-wise” one-to-one comparisons. The relative scores are put into a matrix and the eigenvectors of each row of the matrix are used to calculate a final value for the relative importance of each desired goal. The possible investments or strategic options are then scored against each other for each desired capability, using the process described above. A matrix of the capability scores of each option is used for a final score.

Analysis. ExpertChoice’s greatest advantage could be its mathematically rigorous approach to controlling for variations in the subjective judgments made by ROI software packages—for example, when ranking numerically how important “net thrust” is to a jet engine, or “muzzle velocity” to a ballistic projectile weapon. The relatively large number of individual calculations allows ExpertChoice to create an “Inconsistency Ratio” that is intended to alert the user if the subjective scores vary too widely, although the baseline used is not clear. Unfortunately, the process seems to sacrifice clarity and ease of use for its greater complexity, while the potential accuracy gains are unclear, and possibly not significant. ExpertChoice is structured to consider a variety of alternatives simultaneously, and in the process compresses or displays too awkwardly intermediate data about a single option’s performance against a single objective. Where PALMA allows the user to view an individual tree hierarchy for the performance of every individual investment option *or* a combined graph, ExpertChoice seems to blend its hierarchies together into output tables, and the values are less transparent in origin. Also, ExpertChoice appears no better than PALMA at modeling dynamic interactivity between same-level node performances. Nor does it offer a multi-year data series to facilitate long-term control and evaluation. While this method may be, mathematically, a promising methodology, it is less compatible with our model than PALMA. However, its adoption would by no means be impossible.

Equity is a software platform developed by Catalyze Ltd., based on a methodology developed Professor Larry Phillips for the London School of Economics. The software uses “Multi-Criteria-Decision-Analysis” to analyze the benefits of competing investment or strategy options. This is a simple method: like PALMA, Equity asks the user to create a list of performance objectives and a list of investment options. The user is asked to score the performance of each option against each objective, and the values are stored in a graphical table. Performance objectives can be weighted for their relative importance, and a multiplier can be introduced representing certainty or probability of success. All in all, the methodology is similar to PALMA, except that the hierarchy is presented in the literature as explicitly two-layer. Equity does not appear to use sub-objectives, and does not appear to present a way to process a performance score through multiple rounds of weighted averaging. Like PALMA, Equity can present the overall scores of its performance options on a ‘production frontier’ graph, and walks the user through the same evaluation process in selecting the best options from the graph.

Analysis. Equity can be described with fair accuracy as a cruder version of PALMA. It lacks some of PALMA’s greatest strengths, such as the ability to isolate and visually toggle between a given alternative option and the ‘baseline’ system performance. Equity appears to lack any process at all for constructing hierarchies with more than two layers, while our MOP/MOE/MOV methodology presupposes at least three layers. The lack of layers encourages subjectivity by requiring users to select values for the more conceptual or complex metrics, rather than allowing them to be compiled averages of more empirical metrics. While Equity does allow the forced exclusion or inclusion of certain investments, it does not appear to allow the user to restrict cost to different usage pools, nor to plot performance impact over multiple time increments. It thus underperforms PALMA in *Control* and *Evaluate* stages, while its poor visualizations hinder the *Analysis* stage, forcing oversimplification. Needless to say, the interdependence issues unaddressed by PALMA are not confronted. Equity is an inferior choice for the study.

Balanced Scorecard is not the name of a single software platform, but rather the name of a methodology developed in 1987 and publicized by Lawrence Kaplan of Harvard Business School, who published a series of articles and eventually a book on the subject in the 1990's. "Balanced Scorecard" uses a significantly simpler and less exact methodology than the other software choices previously discussed. Software platforms based on the "Balanced Scorecard" concept are numerous and may vary in individual execution. The core of Balanced Scorecard is a four-box grid, where each box is used to represent a performance category important to the system or organization. The user then fills in each box with metrics believed to be important to the category. This is basically similar to, although less hierarchical than, the *Analysis* stage of our methodology, and a similar process constitutes the first stage of the other software platforms. This is apparently the end of the Balanced Scorecard process, other than drawing 'linkages' between metrics believed to be interrelated. The process does not specifically require or model the evaluation of performance strategies against these metrics. No scores of competing options are created, as there is no requirement for options to be scored.

Analysis. Balanced Scorecard does not qualify as a methodology and is best considered more of a conceptual framework for strategic planning. While there appear to be self-titled "Balanced Scorecard" software packages that implement the recording of performance data against specific goals for the created metrics, there does not appear to be no one established methodology for performing this process.

This completes the discussion of the methodology. The next chapter will briefly discuss how this methodology will be evaluated using an Army S&T area.

Chapter 7. Demonstrating the Methodology

In the next phase of this CTNSP study, the methodology described in this paper will be demonstrated in a use case analysis. The use case will focus on an Army technology area that offers significant future capabilities for our Army, as well as a set of very robust present-day technical challenges. The technology area that appears to be of most interest is robotic platforms. In particular, the methodology will be applied to an Army Research Laboratory program—the Micro-Autonomous Systems and Technology (MAST) Collaborative Technology Alliance (CTA). This will be an exceptionally challenging analysis, since much of the MAST effort focuses on basic research efforts.

The use case will address each of the components of the methodology, apply them to selected current and planned MAST CTA projects, and take into account various topics, capabilities, constraints, and trade-offs—some of which are described below. In addition to demonstrating the applicability of the methodology, the analysis of the MAST CTA should provide its Director useful information for developing an investment strategy.

The remainder of this chapter describes the importance of this technology area, the objective of the CTA, and identifies functional and operational capabilities and constraints of micro-autonomous systems that will be considered in each of the components of the methodology. These capabilities and constraints will be critical to, for example, the decomposition of capabilities into MOPs and MOEs for the use case analysis.

Importance of Robotic Systems

Operations Enduring Freedom and Iraqi Freedom have demonstrated the value of robotic platforms, both aerial and ground, that are teleoperated remotely. Robotic platforms extend the warfighter's senses and reach; they have been used as sensors, communication devices, and, in some instances, weapons platforms. Especially in complex terrain, like caves and mountains, or an urban environment, these platforms provide operational capabilities to the warfighter that would otherwise be costly, impossible, or deadly to achieve with humans. Future enhancements to warfighting capabilities require a reduction in platform size and the cohesive operation of multiple platforms that perform with little or no direct human supervision and can support personnel operating in a variety of dangerous environments.

The CTA

To address these future capabilities for autonomous platforms, in February 2008 the Army established the MAST CTA at its corporate laboratory, the US Army Research Laboratory. The objective of the MAST CTA is to enhance tactical situational awareness in urban and complex terrain by enabling the autonomous operation of a collaborative ensemble of multifunctional, mobile microsystems.

One goal of the MAST CTA is to advance fundamental S&T in several key areas of robotics, including small-scale aeromechanics and ambulation; propulsion; sensing, processing, and communications; navigation and control; microdevices and integration; platform packaging; and systems architectures. The approach is to consider the interplay between all elements, as opposed to each element independently.

To achieve this goal, radical design and engineering methodologies are required in which system-level performance, maneuvering, and functional adaptability are emphasized over the optimization of individual functions. The methodology to be developed may therefore be a critical tool to guide the design of actual platforms.

Functional and Operational Capabilities and Constraints

The impact and interplay between conflicting requirements on the technical issues listed above are so complex that investigating a single issue in isolation of the others will not generate an efficient and operationally effective ensemble of microsystems. Solutions to processing, communications, and mobility, for example, that are satisfactory for large systems do not scale when platforms are reduced to the size considered in the MAST CTA (i.e., palm-sized or smaller). Consider, for example, that platform size and weight limit the power available over the duration of a mission. The largest percentage of available power is utilized for mobility. The limited power in turn constrains the bandwidth of intra-platform communications, e.g., between sensors and processors and processors and transmitters, as well the bandwidth of inter-platform communications. Limited communications further impact the ability of the microsystem collective to sense, understand, and respond coherently as a group.

As stated, the vision of the MAST CTA is to enhance tactical situational awareness in urban and complex terrain. Performers in the CTA were asked to consider three capabilities required to achieve this in three different scenarios: small unit search of an intact building for potential threats, small unit search of demolished building or a cave for potential threats, and perimeter defense of a special operations unit. For the purposes of this study, we consider only the first scenario and refine some of the parameters of the search.

We assume that the search is room by room, and that at least one platform moves autonomously ahead of the troops from room to room. The platforms must provide the troops an indicator of the presence of a large body (e.g., either an infrared image or a simple graphic indicator). The duration of operations in each room is minutes; to clear a building, a half hour to an hour.

Sensing is thus a critical function for these platforms. However, the fidelity and resolution required to insure an effective mission is unknown. This information is critical to platform design. Sensors that can generate high fidelity, high resolution information require more power than low performance sensors and, as mentioned above, power is a limited resource. Thus, a trade-off exists between sensor performance and platform power that is driven by mission effectiveness.

A use case that addresses this issue is beneficial to performers in the MAST CTA and is appropriately scoped to test the methodology presented. As the use case analysis develops, the above capabilities and constraints may be reduced or expanded in quantity and complexity.

Chapter 8. Summary and Concluding Remarks

The Army has always sought the best “bang for the buck” with its S&T investments. However, commercial ROI assessments do not readily apply to Army S&T expenditures. Unlike commercial organizations—where profit, saving money, and market share are important objectives—other factors such as military capability enhancements are key to Army weapon system quality. Army S&T is primarily focused on enhancing capability and reducing technical risk. S&T investments may provide benefit in terms of cost avoidance; not purely to save money, but rather to be able to buy more capability. The Army wants near-term gains to support current operations, but is also very much concerned about enhancing mid- and far-term capabilities. This latter concern complicates the analysis of the benefit of an S&T effort in that the benefit may not be seen for many years, and thus may be affected by the globalization of technology, unforeseen changes in the threat, and other uncontrollable outside influences. To add more complexity to the issue, an S&T effort may morph into other technologies or become a critical component of more than one system.

This complex situation requires a methodology that provides the Army an approach or a few approaches for assessing its S&T investments, though it will still be very challenging to assess non-application-specific basic research efforts. This paper provides such a methodology, dividing it into four main components:

1. “ROI-like” assessments that are focused on typical capability, effectiveness, and performance parameters as well as other areas of benefit (programmatic, life-cycle costs, manufacturability, etc.). Recommendations were also made for using modeling and simulation tools to generate data for providing predictive measures of capability, effectiveness, and performance.
2. Development of a “composite ROI index” that groups many similar benefits together and then can be plotted against funding. Each index would include the sum of multiple terms, with each term having a weight “W,” a requirement “R” (e.g., objective flight duration), and a measure of performance “P” (e.g., actual flight duration). The composite index may be most appropriately used when dealing with numerous parameters, such as those identified in #1 above.
3. Calculation of probability of achieving the measured ROI. This would take into account probabilities of technical success, transition opportunity, technology insertion, adequate funding and ability to maintain a capability advantage.
4. Utilization of a PPM tool to select projects (identified as most promising in the analysis above) for funding based on available budgets, objectives, etc. This paper addressed four PPM tools: PALMA, Balanced Scorecard, Equity, and Expert Choice.

The methodology also incorporates the impact of two growing trends that are already having a significant impact on R&D. One is the topic of globalization, which is summarized in appendix B, but in more detail in separate Defense and Technology Paper.³⁵ Globalization is a two-edged

³⁵ See “Army R&D Collaboration and The Role of Globalization In Research.”

sword. On one hand, it is a poster child for collaboration, bringing together the great minds of the world to cut through a technical barrier. On the other hand, it quickly disseminates the technical information, thus significantly reducing the amount of time that a country may have a technical edge. The methodology attempts to address both the benefits and the problems associated with globalization.

The second trend is internal to the U.S. military—the desire to transform the military into a joint operating force that exploits netcentric capabilities. To support this transformation to “jointness,” netcentric capabilities must be developed which enhance the interoperability of activities, operations, organizations, etc., in which elements of two or more Services participate.

Finally, the paper discusses how, in a follow-on study, we will demonstrate the usefulness of the methodology described herein by applying selected components in a *use case analysis* involving autonomous systems—a technology area that we believe will have high payoff for DOD.

Appendix A. Assessment of UK Approach to Quantifying Benefits of R&D

The Middleton, et al³⁶ paper, was based on a study that examined weapon systems from ten countries—United States, United Kingdom (UK), Russia, France, Japan, Italy, Germany, Australia, Spain, and Sweden. In all, some 69 military systems with the following military service distribution were judged for quality: Army (25 systems), Navy (23 systems), and Air Force (31 systems).

The military system quality was based on an assignment of 1 to 3 points, depending on how a given country's system compared to a like system from another country. Using the main battle tank as an example, tanks from all countries were evaluated and compared with the UK's Challenger 2. In all cases, UK weapon systems were used as the reference and given the following ratings: superior to the UK system (3 points), equal to the UK system (2 points), and inferior to the UK system (1 point). Thus, the total score for the UK was 138 (69 x 2), since it was the country against which all others were graded. All other countries therefore had to lie between 69 and 207, depending on whether their systems were inferior, equal to, or superior to the UK's. Of note is the fact that the US achieved the highest point total (155), while Spain's total (108) was the lowest. These numbers represent an average rating over the time period 1971 to 2005. Then the aggregate scores over all 69 systems were computed for use in the analysis of R&D expenditures.

Over the same time period (1971 to 2005), R&D fiscal data from two reputable sources were collected. The first was from the Organization for Economic Cooperation and Development and the second was from the Stockholm International Peace Research Institute. For the most part, the fiscal data from the different countries were corrected for inflation and then exchange rate adjusted. The resulting R&D expenditure data for the various countries are shown in Table 1.

<i>Country</i>	<i>1951</i>	<i>1956</i>	<i>1961</i>	<i>1966</i>	<i>1971</i>	<i>1976</i>	<i>1981</i>	<i>1986</i>	<i>1991</i>	<i>1996</i>	<i>2001</i>
USA	11698	15533	27114	24498	22138	20250	21685	36977	31889	26685	28546
UK	2491	2636	3346	2927	2590	3295	4034	3315	2866	2510	2170
Russia	3700	4600	13300	21000	27800	35900	44500	53700	47500	1900	3300
France	449	1016	1116	2150	1618	1972	3080	3525	4305	2792	2296
Japan	17	42	70	92	142	132	170	325	504	719	621
Italy	24	33	38	51	66	71	208	401	468	505	263
Germany	11	11	421	765	982	930	785	1166	1267	1116	787
Australia	145	145	143	146	168	155	137	171	154	115	160
Spain	6	6	6	11	8	29	17	55	324	450	1388
Sweden	109	162	175	266	202	230	168	330	397	286	178

Table 1. Government Expenditure on Military R&D (in millions of Pounds Sterling)

³⁶ "The Effect of Defense R&D on Military Equipment Quality," 117-139.

Having the rankings of the country-by-country system quality point totals and the R&D expenditure data in hand, it was possible to statistically correlate the expenditures with system quality points. This was achieved in the study by applying a Spearman analysis to both the expenditure data as well as the system quality points. The result of this analysis was a determination of the degree to which these rankings coincide, by providing a probability measure that the rankings were not random.

By examining the rate with which the system quality improved over time, Middleton et al were able to cross-compare lead times of system quality of one country over another. For example, it was determined that the UK systems had improved at the rate of approximately 3 quality points per year. At this rate it would take about 6 years to reach parity with the US. In this way country-to country lead time comparisons were possible.

It was found that the relationship between lead times and R&D expenditures fit a power law representation. This representation showed that the comparative US position was relatively far out on the lead time versus expenditure curve and was interpreted in the study as possibly being close of the point of diminishing returns. A summary of findings is reported in figure 4.

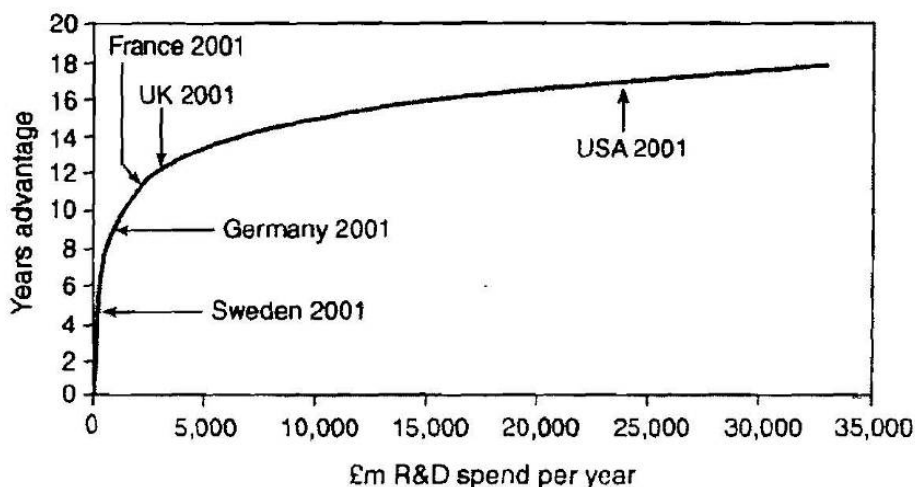


Figure 4. Summary of findings of military expenditures and years of advantage

We feel that such a conclusion may not be warranted given the approach taken by Middleton et al. Their 1-, 2-, and 3-point rating system representing whether or not a given weapon system from the US, Russia, etc is inferior, equal to, or of superior quality to the same UK system lacks fidelity, is very subjective, and limits sensitivity analyses. For example, this rating system would give the same rating to a system that is “slightly better” as it would to a system that is “very much better.” In other words, the process doesn’t go very deeply into quantifying capabilities. Thus, the 3 point gain in quality per nation per year does not appear reassuring, and since this assumption underpins the curve noted above we must express some misgivings regarding conclusions drawn from the appearance of the curve.

Appendix B. Effects of Globalization on Army Research and Development³⁷

The advent of fiber optic communications tying the continents together with very high speed connections, the introduction of ever more powerful computing, the creation of the internet and a host of technologies related to it, and rapid progress in educating scientists and engineers in developing countries—all these are leveling the playing field in world commerce. Friedman coined the phrase “the world is flat” to describe this effect.³⁸ We can see what this has already meant for commerce—the creation of many new multinational corporations; the movement of manufacturing, technical services, and laboratories off shore; and the formation of international coalitions to focus on new opportunities. As an example, when one calls for technical service on a computer, printer, or whatever; the service representative more often than not will be physically located somewhere overseas. India has become adept at capturing these functions by graduating very competent technical people and working very closely and effectively with companies in the United States and elsewhere.

Consider what globalization means in research and development. For many kinds of research it doesn’t matter where the members of a team are located. Given the great bandwidth of communications, they might as well be next door. For many researches, the work can be done entirely on the computer. There appears to be no limit to the size of computer files that can be transferred from one country to another. For theory, for modeling and simulation, and for data reduction and interpretation, distance is no problem. It is not uncommon for a research paper to have authors working continents apart.

There are difficulties in the way of doing Army research across the oceans. Traditionally, collaborative work with other countries has been done by assigning different aspects to one or another partner; for example, for the Abrams tank, the Germans undertook to complete work on the 120mm gun while the Americans fielded the early version with an older 105mm gun perfected by the British. In these cases there was no close day-to-day collaboration. However, recently the Army decided to create a true collaboration with the British on the subject of network and information sciences (the International Technology Alliance on Network and Information Sciences ITA³⁹). Patterned on the Army’s Collaborative Technology Alliances established in the 1990s within the United States, the ITA consists of two consortia, one in the US and one in the UK. The contract document provides the motivation for this arrangement: “The ITA aims to challenge government, industry, and academia to adopt a new way of working. It seeks to break down barriers, build relationships, develop mutual understanding and work in partnership to develop technology for the US and UK military.”⁴⁰ There is joint funding, planning, and execution with exchanges of staff

³⁷See “Army R&D Collaboration and The Role of Globalization In Research.”

³⁸ Thomas L. Friedman, *The World is Flat* (Updated and Expanded), (New York, NY: Farrar, Strauss, and Giroux, 2006).

³⁹ T. Killion, P. Sutton, M. Frame, P. Gendason, “A New Paradigm in International Collaboration: The US-UK International Technology Alliance in Network and Information Sciences,” *RUSI Defense Systems*, June 2007, 46.

⁴⁰ “Other Transaction between the ITA Consortium and the US Army Research Laboratory and the UK Ministry of Defence Concerning Network and Information Sciences,” Agreement No.: 4911NF-06-50001, May 12, 2006, 6.

The advantages of such an arrangement are that it draws on a broader pool of highly trained researchers, shares the costs, and should produce a more capable product. One can imagine such collaborations with our other traditional allies as well as with countries such as India, China, Singapore and the like.

The question is whether such collaborations can become, for the Army, commonplace. Clearly the Army could broaden its horizons and provide the soldier with more and better capabilities. If this is to be the way of the future, the Army will need a cadre of personnel skilled in putting together and managing such arrangements. The challenge is how to deal with security problems. They may not be insurmountable in the basic research arena, but there will be difficulties when moving to applied research and beyond.

Appendix C. “Measuring Return on Investment for Army Basic Research”⁴¹

Introduction

The purpose of this paper is to provide a perspective on the Army’s return on investment for its Basic Research Program. That is, what value does the Army receive for the resources invested in basic research? Experience has shown that the return on investment is significant, both in fiscal terms and in terms of enhancements to the Army’s operational capabilities.

The Army must invest in basic research in order to foster progress in niche areas (such as armor/anti-armor) and where the commercial incentive to invest is lacking due to limited markets (e.g., military medicine to develop vaccines for tropical diseases). In addition, our investment allows us to shape research in other areas to focus on Army-unique issues. This is critical because the fundamental issues that need to be addressed can change when one is focusing, for example, on electrochemistry to support battery power densities applicable to the soldier or on smart materials for real-time management of rotorblade structures as opposed to other applications. Our goal is to conduct focused research to develop and tailor knowledge and technologies to address *Army* needs, to support a full spectrum warfighting force capable of meeting multitude of future mission requirements. This includes both evolutionary research that advances the state-of-the-art and revolutionary concepts that provide the basis for breakthrough capabilities.

However, we also recognize the unpredictable directions that technology may take and the unforeseen benefits that can accrue. When they were first developed, no one would have predicted the widespread applications of lasers, transistors, and fiber optics. We must retain the flexibility to take advantage of technological opportunities as they appear, and to foster their application to meet Army needs. The application process itself can take time as well. Although the transistor, invented in the late 40’s, saw some application in the 50’s in early supercomputers, its wide scale application awaited the development of advanced design and fabrication methods and techniques. Superconductivity was discovered in 1911, and its wide scale application is still awaiting the development of superconducting materials that can operate at reasonable temperatures. The recent discovery of new forms of carbon such as “Fullerene” and “Buckytubes” holds the potential for innovations in areas such as energetic materials and lightweight, high strength fiber. Productive application of such technologies depends on long-term, sustained investment to reap the benefits of innovation.

In assessing our return on investment, or ROI, there are a variety of approaches that we can and should employ. Several methods are highlighted in this paper. The first is a calculation of the benefit-to-cost ratio (BCR). The BCR is calculated by dividing the rate of return per year by the total investment. This approach is used to determine the rate at which one can expect to recover their initial research investment. As an example, a BCR of 0.5 would mean that one’s investment

⁴¹ Thomas Killion and Richard Chait, unpublished paper, written in 1994 while working at Headquarters, Department of the Army, Washington, DC.

is recovered over a two year period. Since the Army does not “profit” from the products of basic research (at least in monetary terms), we measure fiscal benefit in terms of projected cost savings/avoidance instead (e.g., in terms of reduced operating costs). In doing this, we must keep in mind that this is a conservative measure that does not account for additional, non-fiscal benefits such as enhanced military capabilities.

A second approach involves the “Jimmy Stewart Test.”⁴² That is, just as in the movie *It’s a Wonderful Life*, where Jimmy Stewart as George Bailey had the opportunity to witness what would have happened if he had not existed, what capabilities would we *not* have if the Army’s basic research investment had not been made? In particular, what difference have the products of these investments made in our current (or future) military capabilities?

Thirdly, we also consider key scientific measures of merit, which provide an indicator of the impact that our research investment is making within the broader scientific community at large. Measures such as bibliometrics and patents are just two examples of key indicators which can be benchmarked via comparisons with industry or other Government agencies.

Benefit-to-Cost Ratio

There are a multitude of examples from the biomedical research domain where basic research results have led to products which result in significant cost avoidance/savings for health treatment. This includes reduction of risk of disease or injury, reduced treatment time/cost, increased recovery rates, and so on. Table 2 provides specific examples of the medical discoveries fostered by Army research that have led to significant cost avoidance/savings.

Infectious Disease Research	Combat Casualty Care Research	Operational Medicine Research
▪ Malaria prophylaxis ▪ Mefloquine ▪ Adenovirus vaccine ▪ Meningococcus vaccine ▪ Doxycycline ▪ AIDS screening ▪ Arenavirus vaccine ▪ Tularemia vaccine ▪ Alphavirus vaccines ▪ Hepatitis A vaccine ▪ Q-fever vaccine	▪ Advances in vascular surgery and new materials for vascular grafts ▪ Improved protection ▪ Improved early resuscitation & treatment ▪ Bioabsorbable scaffolds for cartilage & tendon repair ▪ Hypertonic saline/dextran resuscitation solution ▪ Intraosseous infusion device ▪ Therapy against nerve agents ▪ Anthrax vaccine ▪ Microencapsulated antibiotics ▪ 6-week blood storage	▪ Epinephrine therapy for glaucoma ▪ Physiological optics advances ▪ Health hazard assessments ▪ Hazard protection ▪ Enhanced soldier performance ▪ Improved treatment of neuropsychiatric injuries ▪ Anti-lipid A monoclonal antibodies

Table 2. Products of Army Medical Research

⁴² Walter L. Robb, “How Good Is Our Research?,” *Research-Technology Management* (March-April 1991), 16–21.

Overall, the estimated cost avoidance/savings of the products listed here is approximately \$1.1 billion per year.⁴³ These technologies are primarily the products of basic research conducted over two decades, from the early 1960's to the early 1980's. If one assumes that these technologies are the result of approximately 20 years of basic research investment across the areas of infectious disease, operational medicine, and combat casualty care, this results in an approximate investment of \$500 million, since the overall investment in medical basic research is approximately \$25 million per year (in today's dollars).⁴⁴ This results in an overall benefit-to-cost ratio (BCR) *greater than 2 to 1*.

In the area of biology, intrinsic chemical markers (ICMs) provide a means for verifying that food is fully cooked, which is essential to destroy the pathogen *C. botulinum* and to be able to confirm that the organism has been destroyed. ICMs allow the use of new thermoprocessing techniques, such as ohmic and microwave cooking, that are more energy efficient and produce a higher quality (i.e., not overcooked) product. By avoiding overprocessing, there is a savings in the energy used to achieve sterilization, and by improving quality there is a savings associated with reduction in wasted or discarded products. These savings translate into lower life-cycle costs for Meals Ready-to-Eat (MREs), Tray-packs, and packaged ration components. The annual savings and BCR can be estimated by considering just the Tray-packs alone. Since there are 600,000 trays purchased annually at a cost of \$18 each, the total annual cost is \$10.8 million. Energy for cooking represents about one-third of the cost and about one-half of the current energy would be required using newer cooking methods with ICMs to verify sterility, resulting in an energy savings of ~\$1.7 million per year. Since the current wastage due to less than optimal quality is about 20%, the savings due to avoiding wastage is about \$2.2 million per year. The total annual savings would be \$3.9 million, based on an initial basic research investment of \$210,000; resulting in a BCR of approximately *19 to 1*.

In the materials research arena, a coating system developed by researchers sponsored by the Army Research Office and the Belvoir Research, Development and Engineering Center greatly reduces corrosion of steel surfaces. Essentially, small amounts of a water-soluble polymer, polyacrylic acid, when added to a zinc phosphate conversion coating solution, control the morphology of the crystallites making up the conversion coating. This results in a product with a 10-fold increase in adhesion, 2-fold increase in flexural properties, and 18+ times improvement in corrosion resistance. The savings to the Army alone in terms of reduced treatment, repair and replacement costs is estimated at over \$600 million/year. Even if one assumes that no other product resulted from the Army materials research program during the period when this product was developed (1984–1987), this results in a BCR of approximately *11 to 1* based on the overall investment in materials research during this period (~\$56 million).

Basic research conducted by the Corps of Engineers on sensor phenomenology and contaminant fate and transport has led to the development of a suite of unique sensors that are deployed on the Site Characterization and Analysis Penetrometer System (SCAPS). Site characterization and analysis is a critical component of the Army's environmental cleanup initiative, which will eventually remediate over 1500 sites. The Corps has documented savings in time and cost of site

⁴³ U.S. Army Medical Research and Development Command, *Medical Science and Technology Master Plan* (Fort Detrick, MD: USARMDC, September 1992).

⁴⁴ All references in this paper to "today's dollars" are 1994 dollars – the year the paper was written.

characterization for one small site, with a savings of over \$300,000 achieved. At a savings of approximately \$500,000 per site, and assuming assessment of 50 sites per year, this results in an annual savings of approximately \$25 million. When compared to the basic research investment of \$1.4 million, this results in a BCR of approximately *18 to 1*.

Two additional Corps of Engineers research products are worthy of mention. In the first one, basic research on thermally stratified flows in a single chilled water storage tank has led to the design for a system that has been adopted for use in the Army and in industry. A demonstration of this system at Fort Jackson led to reduced electrical demand from on-peak to off-peak hours, reducing the fort's electrical bill by \$400,000 annually. Assuming the application of this system to 10 installations, with a approximately \$300,000 savings apiece, this leads to an annual savings of \$3 million and a BCR of approximately *22 to 1*, given the \$135,000 research investment.

The second example involves a "fingerprinting" technique for pavement joint sealants. Premature failure of joint sealants in rigid concrete pavements is typically due to either contractor error (use of wrong materials, improper emplacement, contaminated materials) or government error (wrong specifications). Army and contractor disputes concerning such failure frequently result in litigation. Prior to development of a fingerprinting technique using Fourier Transform Infrared spectroscopy to determine the material properties of the filed sealant, such litigation usually resulted in awards to contractors. The technique has been successfully employed in litigation by the government based on a scientific determination of the cause of sealant failure. The fingerprinting technique either allow s the Government to prevail or deters the vendor from seeking legal recourse, resulting in savings in court costs of approximately \$1 million/year, and a BCR of approximately *3 to 1* based on a basic research investment of \$350,000.

As a final example of calculated BCR, the Army Research Laboratory has developed a mathematical model of the ear which can be used in predicting hearing loss resulting from impulse noise hazards (e.g., rifle, howitzer exposures). The model is currently being considered for adoption as a North Atlantic Treaty Organization (NATO) standard and is being examined by auto manufacturers concerned about airbag design. Use of the model can avert the need for costly facilities and risk to experimental personnel to conduct specialized testing with human subjects to support weapon system design. Comparing the research investment (less than \$2 million) to the cost avoidance (greater than \$1.3 million/year for a specialized facility) results in a BCR of approximately *1 to 1*. This does not include savings in terms of reduced hazards to personnel and increased accuracy in system design and certification.

With the caveat that the focus of our basic research efforts is on enhanced Army effectiveness and added capabilities, the above examples demonstrate that it is possible in some cases to calculate a benefit/cost ratio. Table 3 provides a summary of the calculated BCR for these examples. The BCR ranges from approximately 1 to 1 up to approximately 22 to 1, with a weighted average of about *3 to 1* per year. This represents a conservative estimate, for the reasons cited earlier.

Topic	Calculated BCR
Biomedical research	2 to 1
Combat rations research	19 to 1
Materials research	11 to 1
Environmental research	18 to 1
Chilled water storage	22 to 1
Finger-printing pavement joints	3 to 1
Ear model	1 to 1
WEIGHTED AVERAGE:	3 to 1

Table 3. Calculated BCR for Selected Examples

By way of comparison, representative BCRs from industry are in the range of 0.3 - 0.8 to 1. However, this figure represents BCRs for overall research investment, not just basic research. A reasonable comparison can best be obtained by considering biomedical research. Earlier, we assumed approximately 20 years of basic research investment in the products listed (a total of \$500 million). If we add in 20 years of applied research (a total of \$1200 million @ \$60 million/year) and 20 years of advanced development (\$200 million @ \$10 million/year), this results in a total S&T investment of ~\$1.9 billion. When compared to the overall savings of \$1.1 billion/year, this results in a BCR of approximately *0.6 to 1*. This overall BCR is clearly comparable to results from industry.

More recent BCR data obtained from the Army Natick Soldier Research, Development and Engineering Center (NSRDEC) supports these findings.⁴⁵ The data stem from research for the First Strike Ration (FSR), a compact, eat-on-the-move assault ration intended to be consumed during the first 72 hours of high-intensity conflict. Analysis based on the basic research that was conducted and a cost comparison between the FSR and a comparable MRE yields a BCR of about 1.8 to 1. Considering the entire R&D investment lowers the BCR to about 0.4 to 1. Both values are in line with the earlier BCR estimates noted above.

The “Jimmy Stewart” Test

Our second set of examples is selected based on the “Jimmy Stewart Test.” That is, Army Basic Research was critical to the development of these capabilities. As an example of capability resulting from previous investments, basic research in the early 1980s in the areas of energetic materials, penetration mechanics, and the mechanics of composites, led to the development of the M829A1 for the M1 tank, known as the “Silver Bullet.” The effectiveness of the M829A1 was one of the factors in the overmatching capability displayed by the M1 in Desert Storm. As compared to the prior generation of penetrators, the M829A1 provides a 16% increase in probability of kill (P_K) for the type of threat present in Desert Storm. If an off-the-shelf armor

⁴⁵ John Gassner, e-mail to Richard Chait, April 4, 2008.

appliqué is added to this threat system, this advantage increases to 46%. It is this type of evolutionary or revolutionary technological enhancement that is the goal of our basic research program, ensuring U.S. Army technological dominance on the future battlefield. That, of course, is the ultimate bottom line for our program.

As another example, products of basic research on skill retention and acquisition by the Army Research Institute for the Behavioral and Social Sciences were applied to Operation Joint Endeavour, providing a simple guide to assist trainers in scheduling refresher training. This ensures that soldiers maintain their proficiency in the critical tasks they must perform. Such research has also resulted in adjustments to Army policy on call-ups of Selected Reserve Augmentees, allowing for immediate mobilization of a larger number of Reserves (the SRAs) and rapid train-up.

The Corps of Engineers has developed a multispectral camouflage appliqué (MCA) that combines selectable visual characteristics with high infrared reflectance to produce a truly multispectral camouflage material that can be rapidly applied to existing structures and systems. The versatility of this technology has been demonstrated by its use as a multispectral battlefield marker panel in Bosnia. Versions of this MCA panel are mounted on all NATO tactical vehicles for passive infrared identification. The Joint Combat Identification Office has designated the 4X4 foot version of the MCA panel as the Non-NATO Standard Multispectral Marker Panel for issue to other-than-NATO troops supporting Bosnia operations. It has also been successfully tested and used by the Red Cross. This technology has the potential for saving lives and tens of millions of dollars in fratricide avoidance by aiding accurate identification.

Finally, research at the Natick Research, Development and Engineering Center (NRDEC), part of Soldier and Biological – Chemical Command (SBCCOM), has led to a wide variety of products for the soldier. Basic research with Polaroid led to the development of dyes that absorb laser energy at selected wavelengths, leading to the production of eyewear that protects soldiers from potential blinding. Current research is expanding this protection via the application of dielectric materials that can reflect laser energy at multiple wavelengths to defeat tunable lasers. In another effort, basic researchers developed a model of parachute performance that was applied to identify alternatives to cargo parachutes, which were in short supply. The rapid and successful identification of an alternative ensured continued delivery of critical supplies in Operation Provide Promise.

There are a significant number of capabilities that the Army has today that are direct results of our prior investment in basic research. We would not have these capabilities today were it not for the basic research investments made in the past. This includes the special armor on the M1 tank, night vision systems, advanced penetrators, and the recently demonstrated composite armored vehicle. Our current basic research investments are critical to the emergence of capabilities for the future Army as well, particularly the Army After Next. Examples of key capabilities for the future include electric armaments, full spectrum active protection, and enhanced protection for the soldier.

Scientific Measures of Merit

The last category of measures relates to the scientific merit and achievements of the Army Basic Research program. With regard to scientific measures of merit, figure 5 illustrates the number of patents per R&D \$M and figure 6 illustrates the number of refereed papers per R&D \$M at the Army Research Laboratory (ARL). The graphs indicate that ARL is a leader in the rate of patent generation and is certainly competitive with regard to refereed papers, as compared to industry and other government laboratories.

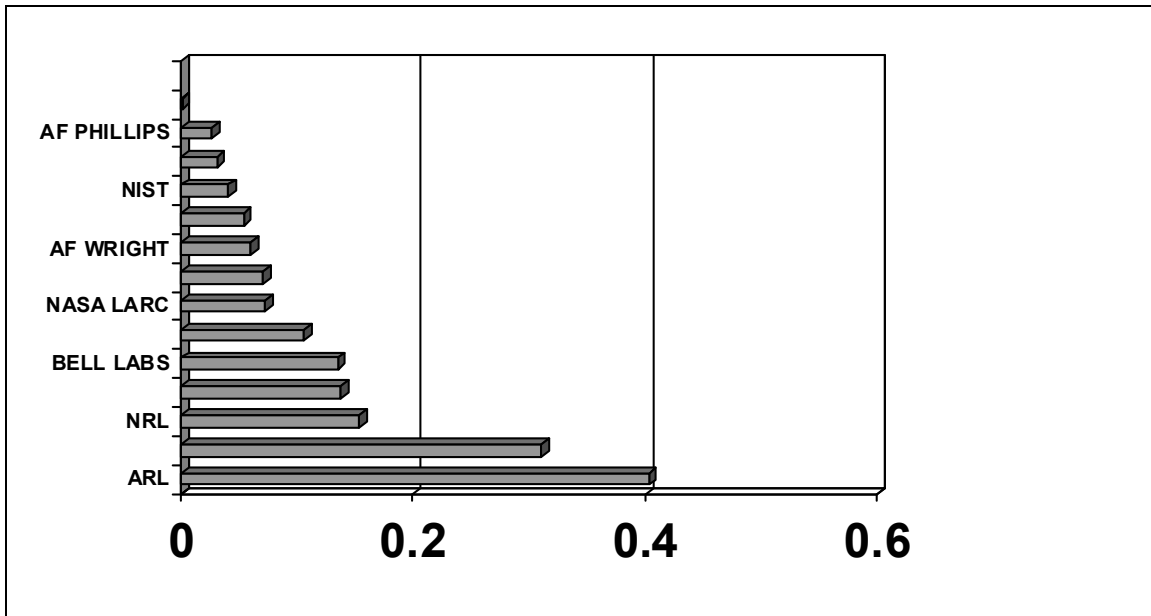


Figure 5. Number of Patents per R&D \$M

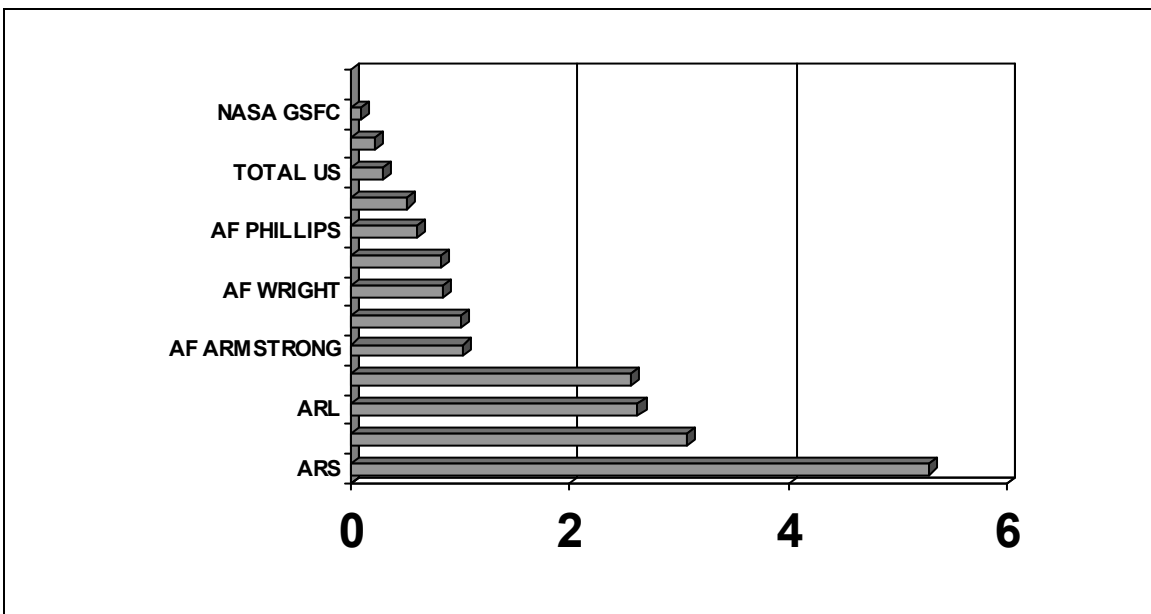


Figure 6. Number of Refereed Papers per R&D &M

Another measure of merit is the quality of the researchers supporting the program. Through the Army Research Office, our extramural program has leveraged world-class researchers across the spectrum of technical areas, including Nobel Prize Winners and innovators in areas such as lasers, superconductors, electronics, and displays, as illustrated in figure 7. These discoveries constitute the central underpinnings for what the National Academy of Engineering has identified as the top ten achievements of the last quarter century.

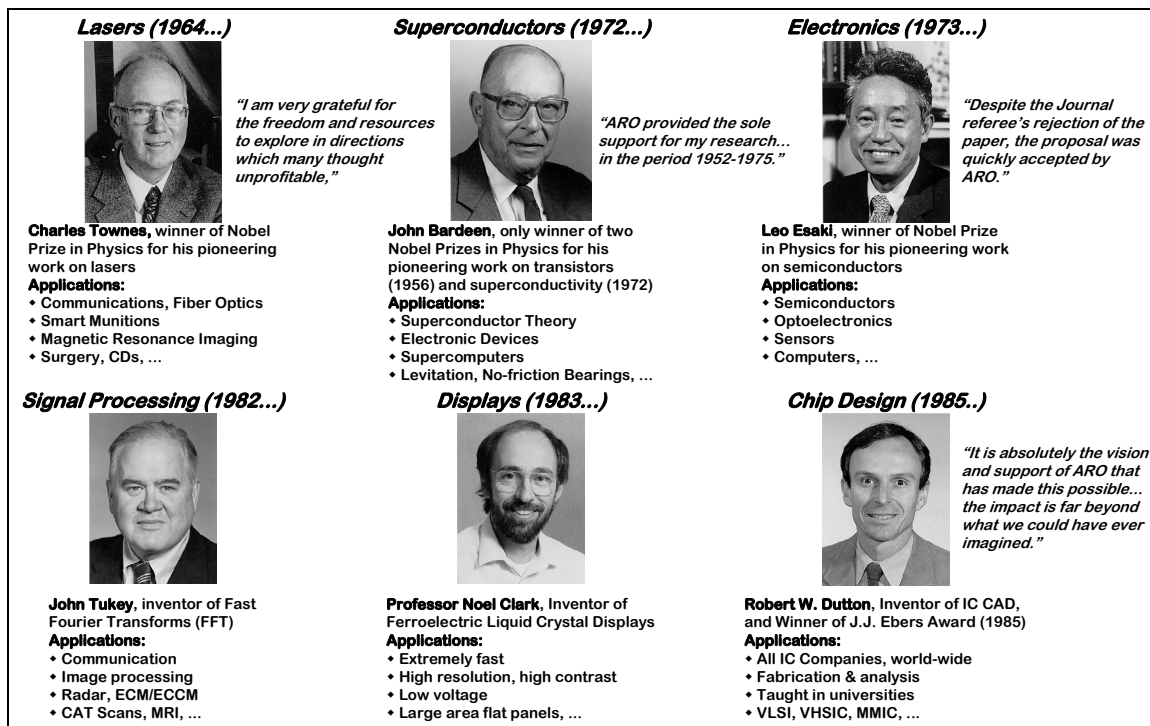


Figure 7. Innovations by ARO-Sponsored Researchers

Future Payoffs

As depicted in figure 8, the Army's Basic Research program has been successful over the years through a multi-faceted approach involving investments in: (1) niche technologies, such as research on missile technologies after World War II that led to the development of the TOW and other missile systems; (2) broader use technologies prior to their commercial viability, such as night vision technologies; and (3) focusing commercial capabilities on Army needs, such as food preparation, packaging, and preservation technologies that enabled the development of D and K rations as well as today's MREs. What can we expect in the future?

As an example, under the Army's Federated Laboratory program, the Army Research Laboratory is exploiting the technical leadership in the private sector to develop advanced display concepts for future command and control systems. Our investment in Display research today will lead to the evolution of new concepts for perceiving, analyzing, and synchronizing the efforts of our battlefield systems, leading to enhanced lethality and survivability for our forces.

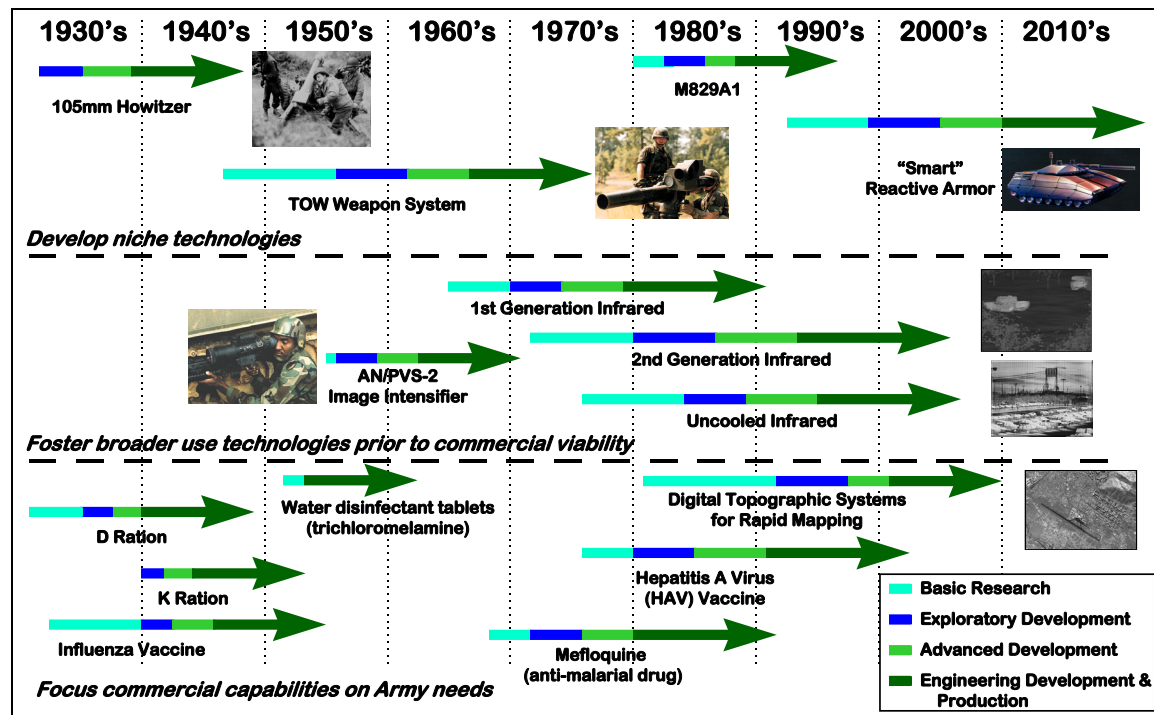


Figure 8. Army Technology Evolution: A Sampler

The Army research community is also fostering the evolution of compact, lightweight power technology to support the future soldier system as well as communications systems, unattended sensors, and other systems. Current investments in chemistry, materials, and other areas will allow the development of higher power density batteries, portable fuel cells, microturbines, and other technologies. Such developments are critical to further reducing the burden imposed on the soldier, as illustrated in figure 9 for Single Channel Ground-Air Radio System (SINCGARS) batteries. Our investments in this area leverage the investments made by the Defense Advanced Research Projects Agency (DARPA) as well as innovations in the private sector.

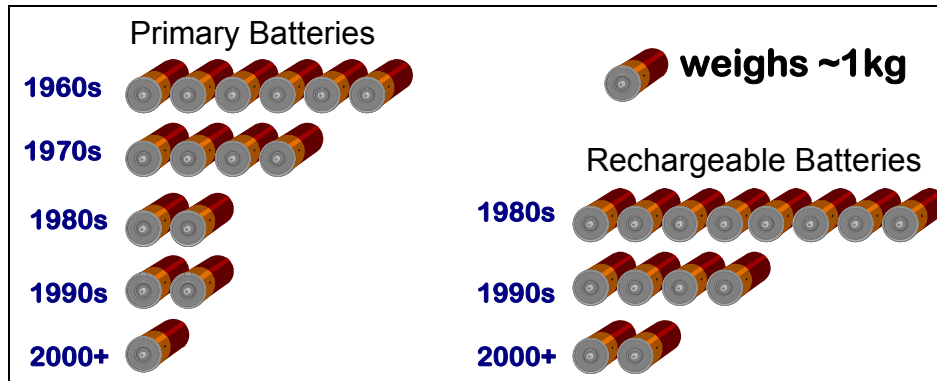


Figure 9. Batteries Required for 48hrs of SINCGARS Operation

Summary

The Army Basic Research program has a history of significant contributions to the Army's capabilities and effectiveness, as well as to the quality of life of its soldiers. This research has been critical to the development of niche technologies that have provided our technological overmatch in armor and anti-armor technologies, as demonstrated in Operation *Desert Storm*. Overall, we believe that these enhanced capabilities speak for themselves in terms of the return on investment for Basic Research. However, where we are able to identify specific cost savings/avoidance resulting from this research, there is an average benefit/cost ratio of 3 to 1 per year, based on conservative estimates. The bottom line is that Army research has made a great difference in the past, and will continue to make a significant difference for the future Army.

Appendix D. Importance of Joint Capabilities in Cost-Benefit Analysis Efforts⁴⁶

What is the value of the Army's S&T investment in interoperable systems, in particular, the investment in network systems that support the functions of battle command and intelligence sharing? Where does S&T investment in interoperability provide the greatest return on investment? And how much interoperability is enough?

The Army invests substantial S&T dollars to make its forces interoperable, most critically through reliable voice and data networks. Interoperability begins inside the Army though the requirement for backward compatibility of new technologies with legacy systems. However, interoperability must also extend to compatibility between the Army and systems of the other Services, allied military forces and—more recently—with civilian agencies of the US government and other civilian entities, especially in post conflict operations.

Dr. Barry's paper attempts to answer the opening questions above, or at least provide a useful framework for their consideration. First, it explores the importance of interoperable capabilities in terms of Army doctrine, that is, how the Army fights and plans to fight in the future. It will examine core and emerging concepts and doctrine to establish the principle that interoperability is a given design feature of any S&T investment and network systems in particular. It will seek to determine high payoff interoperability investment areas, most notably operational network connectivity. It will base its analysis on the three broad categories of interoperability: inter-Service joint interoperability, interoperability in a multinational coalition setting, and interoperability with civil agencies and partners in stabilization and reconstruction operations.

The analysis uses the premise of essential interoperability to propose values in measuring return on interoperability S&T investments. A model is put forth using a priority program in the area of C4 networks to quantify interoperability achieved through investments that target enhancement of interoperable network capabilities. Finally, the analysis proposes a model for visualizing when S&T investment in interoperability begins to show decreasing marginal returns, or the point at which interoperability investment is no longer cost effective. He then discusses how to isolate cost factors related to interoperability and how to aggregate them.

The study concludes with observations and recommendations for further research, including S&T investments not related to networks that will enhance interoperability. S&T managers and decision makers, aware of the essentiality of interoperability as well as the high opportunity costs of designing systems that are not interoperable, can better examine investment proposals in both basic and applied research, and in systems design engineering. They need to know, if the research is successful, will the system play well with others on the Army's core team?

⁴⁶ See "Army S&T Investment Return in Joint, Coalition, and Civil Agency Interoperability."

Appendix E. Development of Metrics

Determining Metrics for an S&T Program

Metrics for an S&T program can be divided into different level of importance—technical, system, and operational. For example, the evaluation of a new obstacle avoidance sensor technology for micro-UAVs may have metrics divided into the three following levels:

Level 1: Technical Performance Metrics. Examples of these metrics are detection range, probability of detection, probability of false alarm, and time needed to process sensor information. These metrics can be called measures of performance (MOPs).

Level 2: System Effectiveness Metrics. System effectiveness metrics may assess the impact of the sensed information on mean acquisition time of obstacles by the micro-UAV, the reaction time of the micro-UAV to miss obstacles, the ability of the military user to “fly” the UAV, and other similar measures. These metrics can be called measures of effectiveness (MOEs).

Level 3: Operational Utility Metrics. Operational utility metrics assess the value of the technology on the ability of the user to sense targets, the staying power of the micro-UAV, the survivability of the micro-UAV, and other operational measures. These metrics can be called measures of value (MOVs).

To ensure that the S&T program metrics are relevant to the S&T program, it would seem prudent to follow a process that:

1. Identifies end-use goals and needs;
2. Derives performance objectives and criterion that relate to these goals/needs (performance objectives are usually expressed in terms of key issues or, as in the Army—essential elements of analysis [EEAs]); and
3. Develops appropriate MOPs, MOEs, and MOVs.

To ensure that there is a close linkage between end-use issues and the selected metrics, the associated objectives and criterion must be carefully mapped to the development of the MOPs, MOEs, and MOVs. There may be many MOPs, MOEs, and MOVs for a particular issue.

So in our case of the obstacle avoidance technology, an MOV, MOE, or MOP is a criterion expressing the extent to which a sensor technology performs a task assigned to the UAV system or technology under a specified set of conditions. An individual MOP, MOE, or MOV supplies a partial answer to a particular question. For example, a MOP would relate to a question like: “How well does Technology X perform when assigned Task Y under a set of conditions Z?” The MOP, MOE, and MOV must be a robust quantitative expression of the degree to which a technology under evaluation meets its objective. Each MOP, MOE, and MOV should have a clearly defined relationship to an end-use goal. Each one should be fully described in an analysis plan that, in turn, provides guidelines for both collecting the appropriate data and avoiding conflicts in data collection.

Technical Approach for Developing Metrics

The technical approach for determining the metrics (MOPs, MOEs, and MOVs) for a technology area can be summarized as follows:

1. Develop analytical requirements for the obstacle avoidance technology and the micro-UAV
2. Determine subjective metrics. Subjective metrics of importance to the system developer and end user and could relate to “operational utility” and are usually non-quantifiable. Some examples of non-quantifiable metrics may be:
 - Usefulness. An end user’s assessment of value added, completeness of information, and accuracy of information are examples of usefulness metrics.
 - Usability. Human factors, interoperability, accessibility, and consistency of information are examples of usability metrics.
 - System Developer’s Assessment of Performance. Standards compliance, overall capability, power consumption, and system availability are examples of metrics that may reflect the system developer’s needs.

Data collection efforts for subjective metrics have to be carefully designed and controlled. One can use online questionnaires, observer notes, video recordings, etc. If possible, subjective metrics must be correlated to quantifiable metrics.

3. Determine Quantifiable Metrics. Quantifiable metrics for obstacle avoidance technologies should:
 - Be relevant to micro-UAV needs (which include system requirements and system specifications)
 - Provide flexibility for identifying new system requirements
 - Be clearly aligned with an objective
 - Be clearly defined, including the data that is to be collected for each metric
 - Identify a clear “cause and effect” or audit trail for the data being collected and the technology being evaluated
 - Minimize the number of variables being measured, with a process identified for deconflicting data collected from and perhaps impacted by more than one variable
 - Identify non-quantifiable effects (e.g., training) and impacts of system “wash-out” (i.e., a technology’s individual performance is lost in the host system performance), and control (or reduce) them as much as possible.

Two caveats about developing quantifiable metrics. First, watch for factors that could affect the validity of the outcome. One could game the outcome or wind up with false/differing results due to conditions set up in the experiment, e.g., the incorporation of weak/improper representations. Second, the need to audit “cause and effect” cannot be overemphasized. With complex system of systems, observed effects are not always self-evident. In developing an audit trail for the data in a complex system of systems, there may be a need for intermediate metrics. There will be at least a need for collecting data at intermediate nodes of a complex system.

4. Document Each Measure. Each MOP, MOE, and MOV description should include at least the following information:

- Title;
- Definition;
- Notation ;
- Limitation/range of the Measure;
- Reasoning;
- Relevance;
- Variations of Use; and
- Associated Measures.

Example of a Documented MOE

For example, here is a relatively simple MOE for an obstacle avoidance technology:

- Title. Mean Acquisition Time of Obstacles
- Definition. Mean acquisition time is the average of elapsed times to complete all successful acquisitions. The acquisition of an obstacle includes the functions of detection, recognition, identification, and location of the obstacle. The elapsed time for each completed acquisition is collected as input data. Thus, mean acquisition time for “n” acquisitions = $(\sum \{\text{time for each successful acquisitions}\} \text{ for } 1 \text{ to “n”}) / \text{“n”}$ successful acquisitions
- Notation. Time in seconds, minutes, or hours (as appropriate).
- Limitation/Range of the Measure. There must be at least 30 successful acquisitions to eliminate differences due to chance deviations. If possible, causes for outlying values should be determined and documented. The output must be a positive number greater than zero.
- Reasoning. This measure addresses timeliness of obstacle acquisition for completed successful acquisitions.
- Relevance. May be used in situations that require obstacle acquisition time as a factor.
- Variations of Use. Can also be defined for “Median Acquisition Time.”
- Associated Measures. Time to detect; time to identify; and expected time to acquire.
Note: Also identify linkages to metrics in ATOs, ATDs, and TTAs.

Appendix F. Acronyms

AHP	Analytical Hierarchical Process
ARCIC	Army Capabilities Integration Center
ARL	Army Research Laboratory
ATD	Army Technology Demonstration
ATO	Army Technology Objective
AWARS	Advanced Warfighting Simulation
BA	Budget Activity
BCR	Benefit-to-Cost Ratio
BCT	Brigade Combat Team
CASTFOREM	Combined Arms and Support Task Force Evaluation Model
CTA	Collaborative Technology Alliance
CTNSP	Center for Technology and National Security Policy
DARPA	Defense Advanced Research Projects Agency
DOE	Design of Experiment
DOTMLPF	Doctrine, Organization, Training, Materiel, Leadership and Education, Personnel, and Facilities
ENIAC	Electronic Numerical Integrator and Computer
FCS	Future Combat Systems
FSR	First Strike Ration
ICM	Intrinsic Chemical Marker
IMPRINT	Improved Personnel Research Integration Tool
JWARS	Joint Warfighting System
LVC	Live, Virtual, and Constructive
MANPRINT	Manpower and Personnel Integration
MAST	Micro-Autonomous Systems and Technology
MATREX	Modeling Architecture for Technology, Research, and Experimentation
MCA	Multispectral Camouflage Appliqué
MIT	Massachusetts Institute of Technology
MOE	Measure of Effectiveness
MOP	Measure of Performance
MOV	Measure of Value
MRE	Meal Ready-to-Eat
NATO	North Atlantic Treaty Organization
NRDEC	Natick Research, Development and Engineering Center
NSRDEC	U.S. Army Natick Soldier Research, Development and Engineering Center
OneSAF	One Semi-Automated Force
PALMA	Portfolio AnaLysis MACHine
PM	Program Manager
PPM	Project Portfolio Management
ROI	Return on Investment
SBCCOM	Soldier and Biological – Chemical Command
SINCGARS	Single Channel Ground-Air Radio System
SoS	System of Systems

SCAPS	Site Characterization and Analysis Penetrometer System
TACWAR	Tactical Warfare
TRAC	TRADOC Analysis Center
TRADOC	U.S. Army Training and Doctrine Command
TRL	Technology Readiness Level
TTA	Technology Transition Agreement
UAV	Unmanned Aerial Vehicle
UK	United Kingdom
VIC	Vector-In-Command