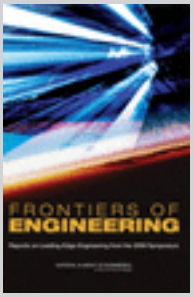


# Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2008 Symposium

National Academy of Engineering

ISBN
978-0-309-12821-6

194 pages
6 x 9
PAPERBACK (2009)

# FRONTIERS OF ENGINEERING

Reports on Leading-Edge Engineering from the 2008 Symposium

NATIONAL ACADEMY OF ENGINEERING
*OF THE NATIONAL ACADEMIES*

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
**www.nap.edu**

**THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, N.W. • Washington, DC 20001**



Funding for the activity that led to this publication was provided by Sandia National Laboratories, The Grainger Foundation, Air Force Office of Scientific Research, Defense Advanced Research Projects Agency, Department of Defense–DDR&E Research, National Science Foundation, Microsoft Research, Sun Microsystems, IBM, Intel, Alcatel-Lucent/ Bell Laboratories, Corning, Inc., Cummins, Inc., and Dr. John A. Armstrong.

International Standard Book Number-13: 978-03-09-12821-6
International Standard Book Number-10: 03-09-12821-8

Additional copies of this report are available from The National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20001; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.



Printed in the United States of America

# THE NATIONAL ACADEMIES

*Advisers to the Nation on Science, Engineering, and Medicine*

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

**www.national-academies.org**

## ORGANIZING COMMITTEE



### Staff

# Preface

In 1995, the National Academy of Engineering (NAE) initiated the Frontiers of Engineering Program, which brings together about 100 young engineering leaders at annual symposia to learn about cutting-edge research and technical work in a variety of engineering fields. The 2008 U.S. Frontiers of Engineering Symposium was hosted by Sandia National Laboratories at the University of New Mexico, September 18-20. Speakers were asked to prepare extended summaries of their presentations, which are reprinted in this volume. The intent of this book is to convey the excitement of this unique meeting and to highlight cutting-edge developments in engineering research and technical work.

## GOALS OF THE FRONTIERS OF ENGINEERING PROGRAM

The practice of engineering is continually changing. Engineers today must be able not only to thrive in an environment of rapid technological change and globalization, but also to work on interdisciplinary teams. Cutting-edge research is being done at the intersections of engineering disciplines, and successful researchers and practitioners must be aware of developments and challenges in areas that may not be familiar to them.

Every year at the U.S. Frontiers of Engineering Symposium, 100 of this country's best and brightest engineers, ages 30 to 45, have an opportunity to learn from their peers about pioneering work being done in many areas of engineering. The symposium gives early career engineers working in academia, industry,

and government in many different engineering disciplines an opportunity to make contacts with and learn from individuals they would not meet in the usual round of professional meetings. This networking may lead to collaborative work and facilitate the transfer of new techniques and approaches. It is hoped that the exchange of information on current developments in many fields of engineering will lead to insights that may be applicable in specific disciplines and thereby build U.S. innovative capacity.

The number of participants at each meeting is limited to 100 to maximize opportunities for interactions and exchanges among the attendees, who are chosen through a competitive nomination and selection process. The topics and speakers for each meeting are selected by an organizing committee of engineers in the same 30- to 45-year-old cohort as the participants. Different topics are covered each year, and, with a few exceptions, different individuals participate.

Speakers describe the challenges they face and communicate the excitement of their work to a technically sophisticated audience with backgrounds in many disciplines. Each speaker provides a brief overview of his/her field of inquiry; defines the frontiers of that field; describes experiments, prototypes, and design studies that have been completed or are in progress, as well as new tools and methodologies, and limitations and controversies; and then summarizes the long-term significance of his/her work.

## THE 2008 SYMPOSIUM

The four general topics covered at the 2008 meeting were: drug delivery systems, emerging nanoelectronic devices, cognitive engineering, and countering the proliferation of weapons of mass destruction. The Drug Delivery Systems session described how advances in materials, particularly polymer systems, have enabled more careful engineering of delivery systems. For example, engineered particles or devices provide sustained release of therapies over an extended time period, eliminating daily dosing, and micro- and nano-engineered systems target delivery of a therapy to a particular physiological system, minimizing systemic side effects. Talks in the session provided an overview of drug delivery methodologies and highlighted several key technologies for targeting and controlling the release of bioactive materials such as targeted polymeric nanotherapies, polymer technology for gene delivery, and traceable drug delivery with quantum dots.

The Emerging Nanoelectronic Devices session focused on novel nanoscale materials and devices, circuit concepts, and sensor functionalities that can be harnessed to develop new technologies for information processing. Presenters discussed a range of ideas for post-CMOS technologies, such as molecular electronics, carbon nanotube devices, and spin devices that when integrated with appropriate nanoarchitectures create alternative electronic devices.

Cognitive Engineering, according to the Human Factors and Ergonomics Society, focuses on improving systems design and training to support human

cognitive and decision making skills, particularly in applied, naturalistic settings. The four presentations in this session provided an overview of the field and described improvements in systems engineering to maximize human performance and reduce error in the domains of driving, power plant operations, and health care delivery.

The symposium concluded with talks on understanding and countering the proliferation of weapons of mass destruction. Two of the talks covered national and international policy issues that frame the discussion, and a third described the strategy of capability-based nuclear deterrence, which relies on a smaller number of deployed weapons and a robust and agile infrastructure enabled by science and engineering.

In addition to the plenary sessions, the participants had many opportunities to engage in informal interactions. On the first afternoon of the meeting, participants broke into small groups to share ideas on important advances they hope to make in the next 10 years and what discoveries would be helpful in reaching their goals. On the second afternoon, there were tours of the Center for High-Technology Materials at the University of New Mexico and two facilities at Sandia National Laboratories: the National Solar Thermal Test Facility (aka the Solar Tower), and the Z Machine, the world's largest X-ray generator.

Every year, a distinguished engineer addresses the participants at dinner on the first evening of the symposium. The speaker this year was Alton D. Romig, Jr., executive vice president and deputy laboratories director for Integrated Technologies and Systems and interim chief operating officer at Sandia National Laboratories. His talk on energy policy and the role of technology in national security covered a range of topics, including engineering advances that provide energy security and the integration and interdependency of world economics and energy markets. The text of Dr. Romig's remarks is included in this volume.

NAE is deeply grateful to the following organizations for their support of the 2008 U.S. Frontiers of Engineering Symposium: Sandia National Laboratories, University of New Mexico School of Engineering, The Grainger Foundation, Air Force Office of Scientific Research, Defense Advanced Research Projects Agency, Department of Defense-DDR&E Research, National Science Foundation, Microsoft Research, Sun Microsystems, IBM, Intel, Alcatel-Lucent/Bell Labs, Corning, Inc., Cummins Inc., and Dr. John A. Armstrong. NAE would also like to thank the members of the Symposium Organizing Committee (p. iv), chaired by Dr. Julia M. Phillips, for planning and organizing the event.

# Contents

**DRUG DELIVERY SYSTEMS**

Introduction 3
*William J. Grieco and Efrosini Kokkoli*

Recent Developments in Needle-Free Drug Delivery 5
*Samir Mitragotri*

Targeted Polymeric Nanotherapeutics 17
*Jeff Hrkach*

Polymer Technology for Gene Therapy 25
*Daniel W. Pack*

Traceable Drug Delivery: Lighting the Way with Qdots 33
*Xiaohu Gao*

**EMERGING NANOELECTRONIC DEVICES**

Introduction 43
*Jia Chen and Victor Zhirnov*

The Quest for the Next Information-Processing Technology 45
*Jeffrey J. Welser*

Molecular and Polymer Nanodevices 53
*Nikolai Zhitenev*

Roll Printing of Crystalline Nanowires for Integrated Electronic and Sensor Arrays 59
*Ali Javey, Zhiyong Fan, Johnny C. Ho, and Roie Yerushalmi*

The Role of DNA in Nanoarchitectonics 71
*Mihrimah Ozkan and Cengiz S. Ozkan*

**COGNITIVE ENGINEERING**

Introduction 85
*Barrett S. Caldwell*

Cognitive Engineering: It's Not What You Think 87
*Stephanie Guerlain*

Driving Attention: Cognitive Engineering in Designing Attractions and Distractions 93
*John D. Lee*

Human Reliability Analysis in Cognitive Engineering and System Design 103
*Ronald Laurids Boring*

Cognitive Engineering Applications in Health Care 111
*Ann M. Bisantz*

**UNDERSTANDING AND COUNTERING THE PROLIFERATION OF WEAPONS OF MASS DESTRUCTION**

Introduction 125
*J. Scott Goldstein and Gregory A. Hebner*

U.S. National Security in New Times 127
*Steven D. Nixon*

Combating Weapons of Mass Destruction: Translating Strategic Guidance into Actionable Solutions 133
*Charles L. Beames*

Nuclear Deterrence in the 21st Century: The Role of Science and Engineering 141
*Joseph C. Martz and Jonathan S. Ventura*

**DINNER SPEECH**

Energy Policy and the Role of Technology in National Security 149
*A.D. Romig, Jr. with Arnold B. Baker*

**APPENDIXES**

Contributors 161

Program 169

Participants 173

# Drug Delivery Systems

# Introduction

WILLIAM J. GRIECO
*PetroAlgae*
*Melbourne, Florida*

EFROSINI KOKKOLI
*University of Minnesota*
*Minneapolis, Minnesota*

Historically, the concept of "drug delivery" has referred to any method of introducing a therapeutic agent into the human body. Traditional delivery methods included oral ingestion, injection, inhalation, and other pathways. Over the past 20 years, advances in materials, particularly polymer systems, have enabled more careful engineering of delivery systems. Current drug-delivery systems now include methods of using traditional mechanisms, such as oral and injection techniques, to introduce engineered particles or devices into the body that can eliminate the necessity of daily doses by providing sustained release therapies. Micro- and nanoengineered systems now offer opportunities to minimize systemic side effects by targeting the delivery of therapies to particular physiological systems.

The presentations in this section provide an overview of drug-delivery methodologies from academic and industrial perspectives. The focus is on polymeric materials for engineering delivery systems. Speakers highlight several key technologies for targeting and controlling the release of bioactive materials.

# Recent Developments in Needle-Free Drug Delivery

Samir Mitragotri
*University of California, Santa Barbara*

Delivering medicines to patients in a safe, effective, and compliant way can be a major challenge (Langer, 2003). Pills and injections are the most common modalities for administering drugs. Although pills can only deliver small molecules, they are generally accepted as a convenient mode of drug delivery (Morishita and Peppas, 2006). Macromolecular drugs such as peptides and proteins, which cannot be taken orally, must be administered by injection. For some drugs, however, systemic administration to healthy tissues can be toxic, regardless of how they are administered. These drugs are only effective if they act directly on specific diseased tissues (Vasir and Labhasetwar, 2005).

The ability of drugs to reach target tissues from the point of administration via pills or injections is limited by the body's multiple barriers, including enzymatic degradation in the stomach, absorption across the intestinal epithelium, hepatic clearance, and accumulation in nontargeted tissues. These barriers have a range of lengths (from the tissue to the organelle level) and time scales.

Collectively, these conditions have made the conversion of potent biomolecules into medical therapies very challenging. The field of drug delivery has

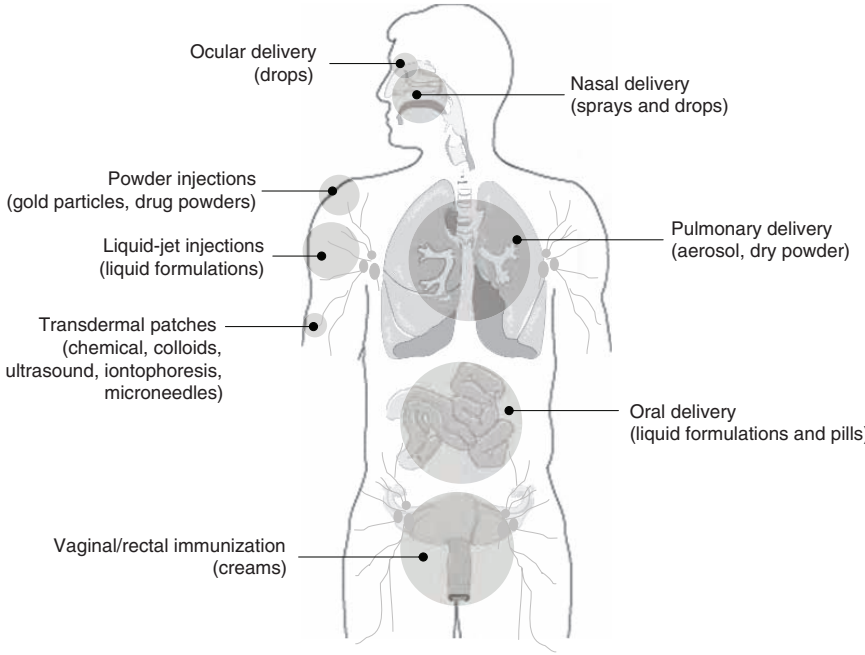


FIGURE 1 Various modes of needle-free drug delivery. Source: Adapted from Mitragotri, 2005.

grown in response to these challenges and is now a significant component of the overall drug development process.

In the past several decades, tremendous progress has been made toward the development of safe, effective, and convenient means of drug administration. Advances have been possible, at least in part, because of our improved understanding of the human body. This article focuses on some key developments in the field of drug delivery, especially those that deal with the development of painless, patient-friendly alternatives to injections for the delivery of macromolecules (Figure 1).

## THE NEED FOR BETTER METHODS OF DRUG DELIVERY

Needles and syringes are the most common method of administering macromolecular drugs; an estimated 12 billion injections are given annually worldwide (Kermode, 2004). Despite their common use, needles have several limitations, including needle phobia (Nir et al., 2003) and accidental needle sticks (Rosenstock, 2000). In addition, concerns have arisen about the unsafe use of needles,

as exemplified by the overwhelming number of HIV, hepatitis C, and hepatitis B infections that are thought to originate each year from the reuse of needles and syringes (Kane et al., 1999).

Noncompliance with medical treatment regimes is also a significant issue. It has been estimated that most patients do not adhere to prescribed dosing regimens, even in developed countries. Noncompliance is linked to several factors, including pain, needle phobia, and forgetfulness, and can result in serious medical complications. In fact, noncompliance is a leading cause of hospitalizations when the carefully designed drug concentration profile is altered in a way that becomes harmful to the patient.

Typically, the blood concentration levels of both injectable and oral drugs that are administered repeatedly vary, depending on the schedule of their administration and the speed at which they are absorbed and distributed by the body. Deviations from the therapeutic range of blood concentrations cause undesirable effects. For these reasons, it is important that drug developers, in addition to considering the efficacy and safety of a drug, must also carefully consider how a drug-delivery system may affect patient compliance.

The limitations of conventional methods of drug delivery can potentially be overcome by needle-free delivery of drugs through the skin or mucosal surfaces of the mouth, nose, or lungs (Varmus et al., 2003). Although these represent viable alternatives to needle-based methods, these surfaces also present significant barriers to drug entry into the body, and breaching them in a safe, effective way is a major goal of drug-delivery research. This article provides a brief review of past efforts, a description of the current status, and prospects for the future, with an emphasis on transdermal and oral drug delivery.

## TRANSDERMAL DRUG DELIVERY

Skin, the largest human organ, provides a painless, compliant interface for systemic drug administration (Zaffaroni, 1991). However, because skin evolved to impede the flux of toxins into the body, it naturally has low permeability to the movement of foreign molecules (Wertz and Downing, 1989). A unique, hierarchical structure of lipid-rich matrix with embedded corneocytes in the stratum corneum (the upper strata [15 μm] of skin) is responsible for this barrier (Wertz and Downing, 1989).

Corneocytes, cross-linked keratin fibers (about 0.2–0.4 μm thick and about 40 μm wide) held together by corneodesmosomes, provide structural stability to the stratum corneum. Lipids, which provide the primary barrier function in the stratum corneum, consist of several components; the primary constituents are ceramides, cholesterol, and fatty acids. The layer of lipids immediately adjacent to the corneocytes is covalently bound to them and plays an important role in maintaining the barrier function. The stratum corneum is continuously desquamated, with a renewal period of about one week, and is actively repaired by the

secretion of lamellar bodies following the disruption of the barrier properties or other environmental insults (Prausnitz et al., 2004).

Transdermal drug delivery involves placing a drug on the skin in the form of a patch, cream, or lotion wherein the drug permeates across the skin and enters the bloodstream. Key advantages of transdermal delivery include the easy accessibility of skin, which encourages patient compliance, avoidance of the gastrointestinal tract, and sustained release over extended periods of time (Prausnitz et al., 2004).

A number of drugs, including scopolamine, nitroglycerin, nicotine, clonidine, fentanyl, estradiol, testosterone, lidocaine, and oxybutinin, are routinely delivered transdermally by skin patches (Prausnitz et al., 2004). The patches, which generally last from one to seven days, depending on the drug, have enabled new therapies and reduced first-pass effects and severe side effects. For example, estradiol patches, which are widely used, have eliminated liver damage, which was a side effect of the drug when it was delivered orally. Transdermal clonidine, nitroglycerin, and fentanyl patches also have fewer adverse effects than the same drugs delivered orally. Nicotine patches have prevented, or at least reduced, smoking and increased lifespans (Prausnitz et al., 2004).

Two classes of transdermal patches are currently available: (1) reservoir-type patches and (2) matrix-type patches. A reservoir-type patch holds the drug in a solution or gel, and the rate of delivery is governed by a rate-controlling membrane. Reservoir-type patches offer more flexibility in terms of drug formulation and tighter control over delivery rates than matrix-type patches. However, they are usually associated with greater design complexity. In matrix-type patches, the drug, adhesive, and mechanical backbone of the patch are combined into a single layer. Thus matrix-type patches are easier to fabricate, but they pose even more significant design constraints than reservoir-type patches (Prausnitz et al., 2004).

Drugs that are currently administered transdermally have two common characteristics—low molecular weight and high lipophilicity. Opening the transdermal route to large hydrophilic drugs, a major challenge in the field of transdermal drug delivery, will require the development of technologies that enable the controlled, reproducible transdermal delivery of macromolecular drugs.

## Passive Methods

Technologies that facilitate transdermal drug delivery can work either passively or actively, depending on whether an external source of energy is used to facilitate skin permeation (Figure 2). Passive methods include chemical enhancers, micelles, liposomes, and peptides (Chen et al., 2006; El Maghraby et al., 2006; Karande et al., 2004; Schreier and Bouwstra, 1994; Schuetz et al., 2005). Examples of chemical enhancers include fatty acids, fatty esters, solvents, and surfactants (Williams and Barry, 1992). These enhancers facilitate transdermal transport by making drugs

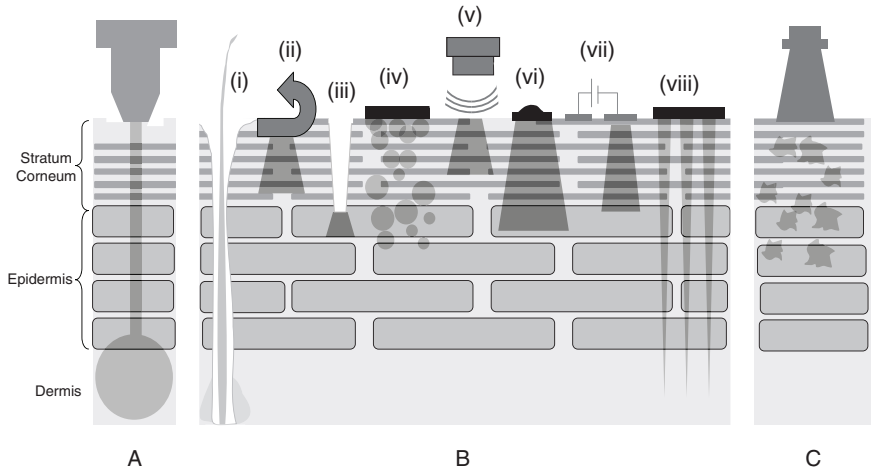


FIGURE 2 Various modes of transdermal drug delivery. (A) Liquid-jet injections deliver drugs into intramuscular, subcutaneous, or intradermal regions. (B) Permeability-based methods of transdermal drug delivery: (i) delivery through hair follicles; (ii) tape-stripping removes the stratum corneum and facilitates drug absorption; (iii) thermal or radio frequency wave-mediated ablation of the stratum corneum creates micropores that enhance drug delivery; (iv) colloidal carriers, such as microemulsions and transfersomes, enhance the dermal absorption of topically applied drugs; (v) low-frequency ultrasound increases drug delivery by making the skin more permeable; (vi) chemical enhancers or peptides for drug delivery; (vii) electroporation of the stratum corneum enhances drug delivery into the epidermis; (viii) microneedles penetrate into the epidermis to deliver drugs. (C) Powder injection delivers dry drug powders into superficial skin layers (epidermis and superficial dermis). Source: Adapted from Mitragotri, 2005.

more soluble, increasing partitioning into the skin, fluidizing the crystalline structure of the topmost layer of skin, or dissolving skin lipids.

Although individual chemical enhancers have had some success, combinations of chemical enhancers are more effective. However, so far, the rational design of combinations of enhancers has been limited by the lack of information on interactions between individual chemical enhancers and the stratum corneum. The number of randomly generated formulations for binary mixtures is in the millions, and the number for higher order formulations (for example, ternary or quaternary mixtures) is even higher. Screening of these formulations is beyond the scope of traditional methods (e.g., Franz diffusion cells).

High-throughput methods of screening transdermal formulations can open this bottleneck and may lead to the discovery of previously unknown mixtures. A new high-throughput method for screening transdermal formulations (Karande et al., 2004) is greater than 100-fold more efficient than Franz diffusion cells (Bro-

naugh, 1989); with this method, up to 1,000 experiments a day can be conducted, an experimental space well beyond the scope of traditional tools (Karande and Mitragotri, 2001). Recent studies have also shown that peptides may effectively increase skin permeability. Specifically, peptides discovered using phage-display methodology have been shown to deliver macromolecules, such as insulin, in vivo (Chen et al., 2006).

Chemical enhancers are relatively easy to incorporate into transdermal patches and can be calibrated to deliver predetermined amounts of a drug by changing the application area. However, passive methods cannot dynamically control the drug dose.

## Active Methods

Active methods can be controlled in real time by varying appropriate parameters. The device and application parameters can also be adjusted to match the patient's skin properties. A growing number of researchers are now exploring transdermal devices with active mechanisms for skin permeation, such as microneedles, jet injectors, ultrasound, iontophoresis, and electrophoresis (Arora et al., 2007; Bashir et al., 2001; Doukas and Kollias, 2004; Habash et al., 2006; Kalia et al., 2004; Karande et al., 2004; Mitragotri et al., 1995; Prausnitz et al., 1993; Zhang et al., 1996).

Microneedles are arrays of micrometer-sized shallow needles that penetrate only into the superficial layers of skin, thereby eliminating the pain associated with standard hypodermic needles (Prausnitz, 2004). Microneedles have been made from a variety of materials, including metals, semiconductors, polymers, and glass, and have been shown to be effective in drug delivery. They have also been produced in solid and hollow forms. Solid microneedles are used to render skin permeable, whereas hollow microneedles actively deliver drugs into the skin at a controlled rate.

In contrast, jet injectors deliver a high-velocity liquid jet stream into the skin, delivering drugs into various skin layers, depending on the jet parameters (Mitragotri, 2006). Jet injectors have a long history, particularly in the delivery of vaccines, insulin, and growth hormone. Ultrasound enhances skin permeability by cavitation, which temporarily disrupts skin structure (Paliwal et al., 2006; Tezel and Mitragotri, 2003). Iontophoresis and electroporation use electric fields to alter skin structure and/or provide additional driving force for drug penetration through the skin (Banga and Prausnitz, 1998; Guy et al., 2000).

## Combined Technologies

Although many individual technologies have been shown to facilitate transderml drug transport, combinations of technologies are often more effective than any of them alone (Mitragotri, 2000). A combination of two or more technologies

may not only increase the enhancement, but may also potentially be safer. Understanding the synergies between technologies and selecting the right combinations is a fruitful area for research that is still largely unexplored.

### Summary

In the last decade, significant new insights have been developed into the structural organization and barrier formation of the skin. In the past, skin was considered primarily a barrier, but it is now known to be a smart material that controls its own structure and function in response to the environment (Menon, 2002). This new knowledge must be incorporated into the future development and evaluation of transdermal technologies.

## ORAL DRUG DELIVERY

Oral drug delivery is the most common, and the preferred type, of drug administration. A large number of small molecules, including those prescribed for the treatment of pain, heart disease, and blood pressure, are already delivered orally. Drugs delivered orally are typically absorbed across the intestinal epithelium into the bloodstream via two mechanisms. The transcellular route involves the transport of drugs through the cell membrane to cross the barrier, either by partitioning of the drug into cell membranes or through the generation of small pores in the outer cell membrane, which allows entry into the cell.

Alternatively, the drug may permeate through the paracellular pathway, which entails transport through the tight junctions between epithelial cells (Cano-Cebrian et al., 2005). A tight junction is a dynamic network of tightly packed proteins in the interstitial spaces of a cell monolayer. Tight junctions have been likened to gatekeepers, as their primary function is to maintain the barrier properties of the epithelium and only permit the transport of very small molecules (< 4 nm in diameter).

A third possibility is that drugs may be actively transported across the epithelium through receptor-mediated endocytosis (Figure 3).

### Proteins and Peptides

The oral delivery of proteins and peptides has elicited a great deal of interest in recent years because of the availability of novel therapeutics through the advent of recombinant DNA technology. Proteins and peptides are macromolecules with a wide variety of functions in biological catalysis, the regulation of cellular processes, and immune-system protection.

Effective oral delivery of a protein or peptide requires that a therapeutic molecule be delivered to the site of interest and cross the intestinal epithelium barrier intact before being transported to the portal circulation system. Unfortunately, this

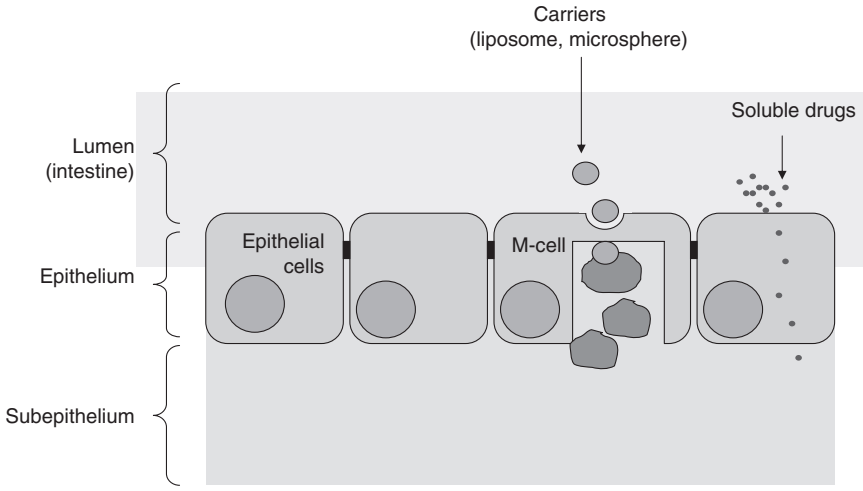


FIGURE 3 Pathways of drug absorption across the intestinal epithelium. Source: Adapted from Mitragotri, 2005.

process is difficult and results in only a small fraction of drug being absorbed in the bloodstream. The delivery of proteins and peptides is further limited by their susceptibility to enzymatic degradation in the gastrointestinal tract (Morishita and Peppas, 2006).

The scientific community has made a major effort in recent years to overcome the obstacles to oral delivery through the development of a large number of new, innovative drug-delivery techniques (Hosny et al., 2002; Luessen et al., 1995; Lyu et al., 2004; Sinha et al., 2004; Whitehead and Mitragotri, 2008; Whitehead et al., 2004, 2008a,b). These methods include enzyme inhibitors, permeation enhancers, mucoadhesive polymers, chemical modification of drugs, targeted delivery, and encapsulation.

## Enzyme Inhibitors

Enzyme inhibitors are used to counteract the natural functions of the enzymes of the gastrointestinal tract that break down ingested proteins. Many studies have been performed in which inhibitors were co-administered with a drug (Bernkop-Schnurch, 1998), but these strategies have seldom been successful unless they included absorption enhancers.

Permeation enhancers have also been used, similar to the way they are used in transdermal drug delivery (Carino and Mathiowitz, 1999). Permeation enhancers, such as surfactants, fatty acids, and bile salts, either disrupt the epithelial

membrane of the intestine or loosen the tight junctions between epithelial cells. While numerous studies have demonstrated that certain enhancers can be very potent delivery aids, safety concerns abound (Aungst, 2000).

### Mucoadhesives

Mucoadhesive strategies have also been used to localize drugs to a small, defined region of the intestine through attractive interactions between the carrier and the intestinal epithelium. This kind of localization results in a high concentration gradient of the drug across the epithelial barrier, which improves drug bioavailability. In addition, a strong adhesion force prolongs the residence time of the dosage at the site of drug absorption, which reduces the dosing frequency and, in turn, increases patient compliance.

Certain mucoadhesive polymers, such as polycarbophil and chitosan derivatives, have been shown to simultaneously act as permeation enhancers and enzyme inhibitors (Luessen et al., 1995; Sinha et al., 2004).

### Encapsulation Technologies

Encapsulation technologies are another alternative for the oral administration of drugs. Using commercially available pH-sensitive polymers, it is possible to target particular regions of the intestine (e.g., jejunum or colon) for drug delivery. Enteric coatings made from these pH-sensitive polymers enable drug-delivery devices to pass through the acidic environment of the stomach unscathed and rapidly dissolve in the intestine. Studies to evaluate these polymers for targeted oral delivery are ongoing in various laboratories (Hosny et al., 2002; Lyu et al., 2004).

Other techniques involve the targeting of M-cells in the intestine to improve mucosal vaccine delivery. M-cells, which are present in the Peyer's patches of the intestine, have the unique ability to take up antigens; targeting can be achieved by using M-cell-specific lectins in combination with a drug-delivery formulation.

Other encapsulation strategies, including microparticles (Mathiowitz et al., 1997), nanoparticles (Carino et al., 2000), and liposomes (Iwanaga et al., 1999), have been developed. These strategies can protect proteins from enzymatic degradation in the intestine and/or facilitate protein uptake across the epithelium (Carino and Mathiowitz, 1999).

## AREAS FOR ONGOING RESEARCH

Novel, painless, patient-friendly methods of drug delivery represent an unmet need in the field of health care. Discoveries in the last decade have demonstrated the feasibility of using several different methodologies for enhancing drug delivery through skin and other mucosal surfaces. These methods have shown the

potential to deliver several molecules, including macromolecules such as insulin and vaccines.

The development of mathematical models to describe and predict transport across the skin and mucosal barriers is another area of active research that has provided useful insights into the development of novel strategies. With the variety of engineering tools at hand, the future of drug delivery looks brighter than ever. The challenge is to convert these discoveries into useful products.

## REFERENCES

Arora, A., I. Hakim, J. Baxter, R. Rathnasingham, R. Srinivasan, D.A. Fletcher, and S. Mitragotri. 2007. Needle-free delivery of macromolecules across the skin by nanoliter-volume pulsed microjets. Proceedings of the National Academy of Sciences of the United States of America 104(11): 4255–4260.

Aungst, B. 2000. Intestinal permeation enhancers. Journal of Pharmacological Science 89(4): 429–442.

Banga, A.K., and M.R. Prausnitz. 1998. Assessing the potential of skin electroporation for the delivery of protein- and gene-based drugs. Trends in Biotechnology 16(10): 408–412.

Bashir, S.L., A.L. Chew, A. Anigbogu, F. Dreher, and H.I. Maibach. 2001. Physical and physiological effects of stratum corneum tape stripping. Skin Research Technology 7(1): 40–48.

Bernkop-Schnurch, A. 1998. The use of inhibitory agents to overcome the enzymatic barrier to perorally administered therapeutic peptides and proteins. Journal of Controlled Release 52(1-2): 1–16.

Bronaugh, R.L. 1989. Determination of Percutaneous Absorption by In Vitro Techniques. Pp. 239–259 in Percutaneous Absorption: Mechanisms, Methodology, and Drug Delivery, edited by R.L. Bronaugh and H.I. Maibach. New York and Basel: Marcel Dekker.

Cano-Cebrian, M.J., T. Zornoza, L. Granero, and A. Polache. 2005. Intestinal absorption enhancement via the paracellular route by fatty acids, chitosans and others: a target for drug delivery. Current Drug Delivery 2(1): 9–22.

Carino, G.P., and E. Mathiowitz. 1999. Oral insulin delivery. Advanced Drug Delivery Reviews 35(2-3): 249–257.

Carino, G.P., J.S. Jacob, and E. Mathiowitz. 2000. Nanosphere based oral insulin delivery. Journal of Controlled Release 65(1-2): 261–269.

Chen, Y., Y. Shen, X. Guo, C. Zhang, W. Yang, M. Ma, S. Liu, M. Zhang, and L.P. Wen. 2006. Transdermal protein delivery by a coadministered peptide identified via phage display. Nature Biotechnology 24(4): 455–460.

Doukas, A.G., and N. Kollias. 2004. Transdermal drug delivery with a pressure wave. Advanced Drug Delivery Reviews 56(5): 559–579.

El Maghraby, G.M., A.C. Williams, and B.W. Barry. 2006. Can drug-bearing liposomes penetrate intact skin? Journal of Pharmacy and Pharmacology 58(4): 415–429.

Guy, R.H., Y.N. Kalia, M.B. Delgado-Charro, V. Merino, A. López, and D. Marro. 2000. Iontophoresis: electrorepulsion and electroosmosis. Journal of Controlled Release 64(1-3): 129–132.

Habash, R.W.Y., R. Bansal, D. Krewski, and H.I. Alhafid. 2006. Thermal therapy, part 1: an introduction to thermal therapy. Critical Reviews in Biomedical Engineering 34(6): 459–489.

Hosny, E.A., H.I. Al-Shora, and M.M. Elmazar. 2002. Oral delivery of insulin from enteric-coated capsules containing sodium salicylate: effect on relative hypoglycemia of diabetic beagle dogs. International Journal of Pharmacology 237(1-2): 71–76.

Iwanaga, K., S. Ono, K. Narioka, M. Kakemi, K. Morimoto, S. Yamashita, Y. Namba, and N. Oku. 1999. Application of surface-coated liposomes for oral delivery of peptide: effects of coating the liposome's surface on the GI transit of insulin. Journal of Pharmacological Sciences 88(2): 248–252.

Kalia, Y.N., A. Naik, J. Garrison, and R.H. Guy. 2004. Iontophoretic drug delivery. Advanced Drug Delivery Reviews 56(5): 619–658.

Kane, A., J. Lloyd, M. Zaffran, L. Simonsen, and M. Kane. 1999. Transmission of hepatitis B, hepatitis C and human immunodeficiency viruses through unsafe injections in the developing world: model-based regional estimates. Bulletin of the World Health Organization 77(10): 801–807.

Karande, P., and S. Mitragotri. 2001. High throughput screening of transdermal formulations. Pharmaceutical Research 19(5): 655–660.

Karande, P., A. Jain, and S. Mitragotri. 2004. Discovery of transdermal penetration enhancers by high-throughput screening. Nature Biotechnology 22(2): 192–197.

Kermode, M. 2004. Unsafe injections in low-income country health settings: need for injection safety promotion to prevent the spread of blood-borne viruses. Health Promotion International 19(1): 95–103.

Langer, R. 2003. Where a pill won't reach. Scientific American 288(4): 50–57.

Luessen, H.I., J.C. Verhoef, G. Borchard, C.M. Lehr, A.G. deBoer, and H.E. Junginger. 1995. Mucoadhesive polymers in peroral peptide drug delivery. II. Carbomer and polycarbophil are potent inhibitors of the intestinal proteolytic enzyme trypsin. Pharmaceutical Research 12(9): 1293–1298.

Lyu, S.Y., Y.J. Kwon, H.J. Joo, and W.B. Park. 2004. Preparation of alginate/chitosan microcapsules and enteric coated granules of mistletoe lectin. Archives of Pharmacal Research 27(1): 118–126.

Mathiowitz, E., J.S. Jacob, Y.S. Jong, G.P. Carino, D.E. Chickering, P. Chaturvedi, C.A. Santos, K. Vijayaraghavan, S. Montgomery, M. Bassett, and C. Morrell. 1997. Biologically erodable microspheres as potential oral drug delivery systems. Nature 386(6623): 410–414.

Menon, G.K. 2002. New insights into skin structure: scratching the surface. Advanced Drug Delivery Reviews 54(Suppl. 1): S3–S17.

Mitragotri, S. 2000. Synergistic effect of enhancers for transdermal drug delivery. Pharmaceutical Research 17(11): 1354–1359.

Mitragotri, S. 2005. Immunization without needles. Nature Reviews. Immunology 5(12): 905–916.

Mitragotri, S. 2006. Current status and future prospects of needle-free liquid jet injectors. Nature Reviews. Drug Discovery 5(7): 543–548.

Mitragotri, S., D. Blankschtein, and R. Langer. 1995. Ultrasound-mediated transdermal protein delivery. Science 269(5255): 850–853.

Morishita, M., and N.A. Peppas. 2006. Is the oral route possible for peptide and protein drug delivery? Drug Discovery Today 11(19-20): 905–910.

Nir, Y., A. Paz, E. Sabo, and J. Potasman. 2003. Fear of injections in young adults: prevalence and associations. American Journal of Tropical Medicine and Hygiene 68(3): 341–344.

Paliwal, S., G.K. Menon, and S. Mitragotri. 2006. Low-frequency sonophoresis: ultrastructural basis for stratum corneum permeability assessed using quantum dots. Journal of Investigative Dermatology 126(5): 1095–1101.

Prausnitz, M.R. 2004. Microneedles for transdermal drug delivery. Advanced Drug Delivery Reviews 56(5): 581–587.

Prausnitz, M.R., V.G. Bose, R. Langer, and J.C. Weaver. 1993. Electroporation of mammalian skin: a mechanism to enhance transdermal drug delivery. Proceedings of the National Academy of Sciences of the United States of America 90(22): 10504–10508.

Prausnitz, M.R., S. Mitragotri, and R. Langer. 2004. Current status and future potential of transdermal drug delivery. Nature Reviews. Drug Discovery 3(2): 115–124.

Rosenstock, L. 2000. Needlestick Injuries Among Healthcare Workers. Statement of the Director, CDC National Institute for Occupational Safety and Health, before the House Subcommittee on Workforce Protections, Committee on Education and the Workforce, U.S. House of Representatives, June 22, 2000.

Schreier, H., and J. Bouwstra. 1994. Liposomes and niosomes as topical drug carriers: dermal and transdermal drug delivery. Journal of Controlled Release 30(1): 1–15.

Schuetz, Y.B., A. Naik, R.H. Guy, and Y.N. Kalia. 2005. Emerging strategies for the transdermal delivery of peptide and protein drugs. Expert Opinion on Drug Delivery 2(3): 533–548.

Sinha, V.R., A.K. Singla, S. Wadhawan, R. Kaushik, R. Kumria, K. Bansal, and S. Dhawan. 2004. Chitosan microspheres as a potential carrier for drugs. International Journal of Pharmaceutics 274(1-2): 1–33.

Tezel, A., and S. Mitragotri. 2003. Interactions of inertial cavitation bubbles with stratum corneum lipid bilayers during low-frequency sonophoresis. Biophysical Journal 85(6): 3502–3512.

Varmus, H., R. Klausner, E. Zerhouni, T. Acharya, A.S. Daar, and P.A. Singer. 2003. Public health. grand challenges in global health. Science 302(5644): 398–399.

Vasir, J.K., and V. Labhasetwar. 2005. Targeted drug delivery in cancer therapy. Technology in Cancer Research & Treatment 4(4): 363–374.

Wertz, P.W., and D.T. Downing. 1989. Stratum Corneum: Biological and Biochemical Considerations. Pp. 1–17 in Transdermal Drug Delivery: Developmental Issues and Research Initiatives, edited by J. Hadgraft and R.H. Guy. New York and Basel: Marcel Dekker.

Whitehead, K., and S. Mitragotri. 2008. Mechanistic analysis of chemical permeation enhancers for oral drug delivery. Pharmaceutical Research 25(6): 1412–1419.

Whitehead, K., Z. Shen, and S. Mitragotri. 2004. Oral delivery of macromolecules using intestinal patches: applications for insulin delivery. Journal of Controlled Release 98(1): 37–45.

Whitehead, K., N. Karr, and S. Mitragotri. 2008a. Discovery of synergistic permeation enhancers for oral drug delivery. Journal of Controlled Release 128(2): 128–133.

Whitehead, K., N. Karr, and S. Mitragotri. 2008b. Safe and effective permeation enhancers for oral drug delivery. Pharmaceutical Research 25(8): 1782–1788.

Williams, A.C., and B.W. Barry. 1992. Skin absorption enhancers. Critical Reviews in Therapeutic Drug Carrier Systems 9(3-4): 305–253.

Zaffaroni, A. 1991. Overview and evolution of therapeutic systems. Annals of the New York Academy of Sciences 618: 405–421.

Zhang, I., K.K. Shung, and D.A. Edwards. 1996. Hydrogels with enhanced mass transfer for transdermal drug delivery. Journal of Pharmaceutical Science 85(12): 1312–1316.

# Targeted Polymeric Nanotherapeutics

JEFF HRKACH
*BIND Biosciences Inc.*
*Cambridge, Massachusetts*

This paper provides an overview of steps being taken by BIND Biosciences Inc. to translate innovative research conducted at the Massachusetts Institute of Technology (MIT) and Harvard Medical School into novel, targeted, polymeric nanotherapeutics.

Advances in drug delivery have significantly affected the lives of patients afflicted with a variety of diseases. New drug-delivery strategies can improve the efficacy, safety, and/or compliance of existing approved medicines and can lead to the development and approval of new drugs with inherent properties (e.g., solubility, bioavailability, off-target side effects) that might otherwise keep them from being approved. In many cases, these improvements are the result of changes in formulation leading to, for example, longer lasting action or a change in delivery modality (e.g., transdermal or inhalation).

Particle-based drug delivery, particularly polymeric particle systems wherein delivery is achieved by encapsulation, or physical entrapment, of a drug within the particle matrix, has been a very active area of interest that has resulted in several successful products. One example is Risperdal CONSTA®, which is indicated for the treatment of schizophrenia.[1]

[1]See *http://www.risperdalconsta.com.*

Risperdal delivers the drug risperidone encapsulated in poly(lactide-co-glycolide) (PLGA) biodegradable polymeric microspheres with a particle diameter of about 100 microns via intramuscular injection once every two weeks. The drug is released over time from the particles by slowly diffusing out of the polymeric matrix as water diffuses in and as the polymer chains degrade via hydrolysis, causing particles to lose their structure and fall apart. PLGA-based particle drug-delivery systems can be tailored to the properties of the drug, appropriate dosage, and the mechanism of action for releasing the encapsulated drug over a period of weeks or months in a controlled way.

Risperdal®, the original risperidone product, is taken orally by patients with schizophrenia on a daily basis. In most cases, the simplicity of taking a pill is very strongly preferred as a method of administering a drug, and designing a drug-delivery system to change the administration from oral delivery to a more complicated (e.g., inhalation) or painful (e.g., injection) delivery, would normally be unsuccessful (unless the oral drug had a significant shortcoming).

For patients with schizophrenia, however, taking a pill every day can be problematic, and missing a dose one day can lead to a downward spiral of missing more doses. In this case, intramuscular injection administered by a doctor or nurse once every two weeks has not only increased patient compliance, but also improved the efficacy of the drug, resulting in a significant improvement in the treatment of patients with schizophrenia.

Microparticle delivery systems, such as Risperdal CONSTA, are too big to be administered intravenously. Their particle size would result in very fast clearance by the body's defense mechanisms or could potentially pose a significant safety risk if they were to lodge in capillary beds in the heart or lungs. Nanoparticle-based drug-delivery systems, in which particle sizes generally range from about 20 to 200 nanometers, are being investigated for delivering therapeutic agents, imaging diseased tissues or organs, and sensing the effectiveness of drug delivery or the status of disease. As a point of reference, a nanometer is one-billionth of a meter or one-millionth of a millimeter. Because of their very small size, nanoparticles administered systemically (i.e., by intravenous injection or infusion) circulate through the bloodstream carrying their therapeutic payloads directly to the site of disease in the body.

## NANOPARTICLE-BASED DRUG-DELIVERY SYSTEMS

Diseases associated with defects or irregularities in the endothelial cells of blood vessels in the diseased area, creating what is called "leaky vasculature," may be particularly susceptible to treatment by nanoparticle-based drug-delivery systems. These include inflammatory diseases (e.g., rheumatoid arthritis, atherosclerosis), infectious diseases (e.g., tuberculosis), and cancer. Once nanoparticles reach the affected area, they can passively diffuse from the bloodstream across the leaky vasculature to deliver drugs directly to the disease site.

However, because nanoparticles are foreign bodies circulating in the bloodstream, the natural defense mechanisms of the body attempt to remove them. The way the body protects itself from nanoparticles or other foreign particulate matter circulating in the bloodstream is through the mononuclear phagocytic system (MPS), sometimes also called the reticuloendothelial system, in which phagocytic cells located primarily in the liver and spleen engulf the nanoparticles. High levels and fast rates of nanoparticle clearance by the MPS lead to an accumulation of nanoparticles in the liver and spleen, thus removing them from circulation before they are able to reach the site of disease and effectively deliver their therapeutic payloads. In addition, if the drug being delivered has potential specific toxicities in the liver or spleen, the clearance of nanoparticles by these organs may exacerbate those effects, making the drug less tolerable or more dangerous.

The optimization of nanoparticle properties, therefore, is critical to the development of a safe nanoparticle drug-delivery system. Particle-surface characteristics (e.g., chemical composition, charge) have a strong influence on the detection of nanoparticles by the MPS. Therefore, one way to minimize MPS clearance is to construct nanoparticles with poly(ethylene glycol) (PEG), a biocompatible polymer, on the surface, a technique that has been successfully used to increase the circulation time of biodegradable polymeric nanoparticles (Gref et al., 1994). The hydrophilic, uncharged nature of PEG can interfere with phagocytic recognition and the uptake of nanoparticles or proteins resulting in prolonged circulation times and more opportunity for the drug to reach the intended disease target.

DOXIL®, a liposomal formulation of the drug doxorubicin that uses a PEG surface to prolong circulation time,[2] is approved for treatment of ovarian cancer, AIDS-related Kaposi's sarcoma, and multiple myeloma. Doxorubicin, like many drugs, does not have a long circulation time in the bloodstream but instead can diffuse throughout the body in a way that can cause untoward side effects and that limits the amount of drug delivered to the tumor, thus decreasing its efficacy. By encapsulating doxorubicin in PEGylated liposome nanoparticles, DOXIL allows for longer circulation times than the drug has in its free, unencapsulated state, in fact long enough for the particles to diffuse into and deliver doxorubicin to the tumor vasculature.

A potential downside of nanoparticle-based drug-delivery systems is that they can deliver more drug to certain parts of the body than the free drug would normally deliver, which can result in either new side effects or an exacerbation of existing side effects. For DOXIL, the result is an increase in the incidence of hand-foot syndrome (a skin irritation that usually occurs on the hands and feet) compared to doxorubicin alone. The apparent cause is that the long-circulating nanoparticles eventually land in the capillary beds of the hands and feet where they deliver liposome-encapsulated doxorubicin in greater amounts than would be delivered by free, unencapsulated doxorubicin.

[2]See *http://www.doxil.com/*.

To repeat, nanoparticles can passively diffuse from the circulating bloodstream through the leaky defects in tumors or areas of infection or inflammation to deliver their therapeutic payloads. Although effective, this passive targeting can have limitations in that nanoparticles may also diffuse out of the disease site through the defects back into circulation. Considerable research is being conducted to improve nanoparticle drug-delivery systems by trying to actively target the nanoparticles to diseased cells (Allen, 2002; Heidel et al., 2007; Peer et al., 2007). These approaches attempt to take advantage of the presence of unique or highly up-regulated cell-surface receptors on diseased cells by functionalizing the surface of nanoparticles with ligands that promote cell-specific recognition and binding.

The intent is that once the particles successfully migrate through the bloodstream to the disease site, targeted nanoparticles will then anchor themselves to the disease cells, keeping the nanoparticles in place long enough to deliver their payloads. The choice and properties of the cell-surface receptor may even allow for the uptake of intact nanoparticles into the cell. The resulting intracellular drug delivery can greatly increase the effectiveness of the drug.

For some drugs and therapeutic applications, intracellular delivery may be necessary, thus requiring intracellular nanoparticle trafficking. One example of this is the new class of short-interfering RNA (siRNA) drugs, which are being developed to inhibit the production of disease-causing proteins through RNA interference (RNAi).

## THE BIND TARGETED NANOPARTICLE

BIND Biosciences Inc. (BIND), a biopharmaceutical company that was founded upon the research of two pioneers in nanoparticle drug delivery, Professor Robert Langer of MIT and Professor Omid Farokhzad of Brigham and Women's Hospital of the Harvard Medical School, has developed methods of engineering targeted nanoparticles composed of biodegradable and biocompatible polymers with precise biophysicochemical properties optimized to deliver drugs for specific therapeutic applications (Gu et al., 2008).

The foundational research by Langer and Farokhzad put BIND in a position to pursue the development of targeted polymeric nanotherapeutics for treating several diseases. BIND's lead program is focused on translating their innovative academic findings into improved treatments for patients with cancer. The BIND technology offers a unique combination of long-circulating nanoparticles with the capability of targeting diseased cells specifically and releasing drugs from nanoparticles in a programmable, controlled way.

Figure 1 is a schematic diagram of a BIND targeted nanoparticle. The targeting ligand enables the nanoparticle to recognize specific proteins or receptors on the surface of cells involved in disease, or in the surrounding extracellular matrix, and bind, with high specificity and avidity, to its intended cellular target

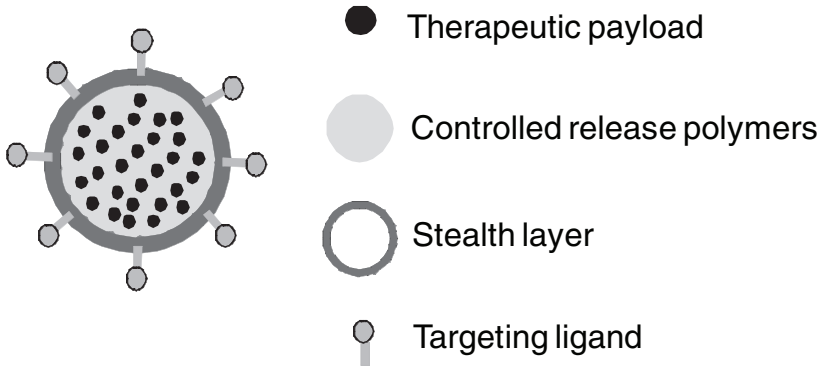


FIGURE 1 Schematic diagram of a BIND targeted polymeric nanoparticle.

site. Many types of cancer have been shown to have cell-surface receptors that are highly expressed on the cancer cells (e.g., prostate cancer [prostate-specific membrane antigen, PSMA], breast cancer [human epidermal growth factor receptor 2, HER-2], and lung cancer [epidermal growth factor receptor, EGFR]), and many drugs are being evaluated that might improve treatment outcomes.

## Surface Functionalization

Surface functionalization imparted by a PEG component shields the targeted nanoparticles from MPS immune clearance, while providing an attachment site for the targeting ligand on the particle surface at precise, controlled levels through proprietary linkage strategies. A key to the successful development of BIND targeted nanoparticles is the optimization of the nanoparticle surface, which requires a precise balance between the targeting ligand and PEG coverage so the nanoparticle surface is masked enough to provide circulation times long enough to reach the disease site and enough targeting ligand on the surface to effectively bind to the target cell-surface receptors. This delicate balance requires precise control over the nanoparticle production process. It also requires the discovery and selection of ligands that are potent and specific enough to bind selectively to the targeted disease cells while remaining bound to the nanoparticle surface.

The polymer matrix, the bulk of the nanoparticle composition, encapsulates the drug in a matrix of clinically safe, validated biodegradable and biocompatible polymers that can be designed to provide appropriate particle size, drug-loading level, drug-release profile, and other critical properties. A variety of drugs or therapeutic payloads can be incorporated into the targeted nanoparticles, including small molecules, peptides, proteins, and nucleic acids, such as siRNA.

### Drug-Release Profile

The drug-release profile is a critical factor for the effective delivery of targeted nanoparticles. If the drug leaks out of the nanoparticle too quickly, it will be released into the bloodstream and essentially delivered as free, unencapsulated drug, thus losing the advantages of nanoparticle delivery. If the drug is not released in the appropriate time frame after the nanoparticles have reached the disease site, it may not reach an efficacious level. Thus it is critical that the right combination of polymer properties be tailored to ensure the optimal drug-release profile. BIND targeted polymeric nanotherapeutics can be engineered with different physicochemical properties, mechanisms of action, and dose requirements to provide effective drug delivery for a variety of diseases with different indications.

## REGULATORY REQUIREMENTS

When a start-up company is founded based on academic research, the initial scientific efforts are focused on transferring the technology from academic laboratories to the company, where researchers can establish the capabilities of the technology and reproduce the results. Shortly thereafter, with a baseline understanding of the technology in hand, the translational aspects of the research begin. The company focuses on defining the most suitable disease indications to pursue and the specific characteristics required.

At this point, the regulatory requirements dictated in the United States by the Food and Drug Administration (FDA) for pharmaceutical development of drug product candidates must be taken into consideration. Since its inception in early 2007, BIND has undertaken a combinatorial optimization approach resulting in a number of enabling improvements to nanoparticle formulation, as well as the nanoparticle production process to meet the needs of its lead targeted oncology candidate.

The optimization approach includes evaluating the performance of nanoparticles using in vitro cell-based assays and in vivo preclinical testing, as well as several chemistry, manufacturing, and controls (CMC) requirements mandated by current manufacturing practices and the FDA to ensure, among other things, batch-to-batch reproducibility and shelf-life stability. Meeting these requirements entails testing a variety of properties, such as particle size, content of the targeting ligand, drug-loading level, and the stability of the nanoparticles and the drug under storage and in-use conditions. As pharmaceutical development progresses, the CMC requirements become more stringent. However, even at this early stage, the company begins testing critical parameters.

To establish an acceptable level of safety and tolerability to support the initial evaluation of a candidate drug product in human clinical studies, the FDA requires formal safety testing in animal models. This is the first major step in the FDA-regulated area of pharmaceutical development. It also represents the

company's first efforts at scaling up the formulation and process capabilities of the drug. Whereas research at MIT/Harvard and initial efforts at BIND were conducted on nanoparticle batches prepared on the bench-top milligram scale, BIND nanoparticle production batch size has been scaled up three orders of magnitude for the animal safety and tolerability tests that support clinical studies.

The critical, long-term stage of pharmaceutical development is clinical testing. Through a progression of studies, the safety, tolerability, and efficacy of a drug product candidate are established; the tests are accompanied by a series of submissions to and discussions with the FDA.

For BIND targeted polymeric nanotherapeutic drug candidates based on improving the performance of existing marketed drugs, the clinical testing period is likely to be shorter than for a completely new drug candidate, because the history and data established for the existing drug provide valuable reference points for BIND and the FDA. Nevertheless, several clinical studies are required, all CMC requirements must be met, and the nanoparticle production process must be scaled up to the kilogram level to supply the drug for clinical studies and ultimately, if successful, to supply the approved, marketed drug to doctors and patients.

Thus a long, challenging, very exciting pathway lies ahead for BIND Biosciences in translating the novel targeted polymeric nanoparticle drug-delivery research by Professors Langer and Farokhzad into medicines that can improve, and even save, the lives of patients suffering from serious diseases.

## REFERENCES

Allen, T.M. 2002. Ligand-targeted therapeutics in anticancer therapy. Nature Reviews 2(10): 750–763.

Gref. R., Y. Minamitake, M.T. Peracchia, V. Trubetskoy, V. Torchilin, and R. Langer. 1994. Biodegradable long-circulating polymeric nanospheres. Science 263(5153): 1600–1603.

Gu, F., L. Zhang, B.A. Teply, N. Mann, A. Wang, A.F. Radovic-Moreno, R. Langer, and O.C. Farokhzad. 2008. Precise engineering of targeted nanoparticles by using self-assembled biointegrated block copolymers. Proceedings of the National Academy of Sciences of the United States of America 105(7): 2586–2591.

Heidel, J.D., Z. Yu, J.Y. Liu, S.M. Rele, Y. Liang, R.K. Zeidan, D.J. Kornbrust, and M.E. Davis. 2007. Administration in non-human primates of escalating intravenous doses of targeted nanoparticles containing ribonucleotide reductase subunit M2 siRNA. Proceedings of the National Academy of Sciences of the United States of America 104(14): 5715–5721.

Peer, D., J.M. Karp, S. Hong, O.C. Farokhzad, R. Margalit, and R. Langer. 2007. Nanocarriers as an emerging platform for cancer therapy. Nature Nanotechnology 2(12): 751–760.

# Polymer Technology for Gene Therapy

Daniel W. Pack
*University of Illinois at Urbana-Champaign*

Gene therapy can be defined as the treatment of human diseases by the transfer of genetic material into specific cells to elicit a desired therapeutic phenotype. It is not difficult to envision treating monogenic diseases, such as hemophilia, muscular dystrophy, and cystic fibrosis, by replacing errant genes within the affected cells. Gene therapies are also being developed, however, for treating cardiovascular, neurological, and infectious diseases, wound healing, and cancer, by delivering genes to augment naturally occurring proteins, to alter the expression of existing genes, or to produce cytotoxic proteins or prodrug-activating enzymes.

Because of the broad potential of gene therapy, it has been heavily investigated during the past 30 years. The first clinical trial of gene therapy, for the treatment of severe combined immunodeficiency (SCID), was initiated in 1990 (Blaese et al., 1995), but it took until April 2000 before the first clinical success was reported by Cavazzana-Calvo et al. (2000) of the treatment of two infants with γc-SCID. Also that year, Kay et al. (2000) reported positive data, including increased circulating levels of factor IX in a hemophilia clinical trial, and Khuri et al. (2000) reported a successful Phase II trial using a combination of gene therapy and traditional chemotherapy to treat recurrent squamous-cell carcinoma

of the head and neck. However, despite more than 1,300 clinical trials to date, no products have been FDA-approved.

At the same time, tragic setbacks, including the deaths of patients in two trials, have hindered progress. A severe inflammatory response caused by the adenovirus used in a 1999 trial for the treatment of ornithine transcarbamylase deficiency was proved to be the cause of death and resulted in a temporary halt to all gene therapy trials. In addition, at least 2 of the 11 children in the Cavazzana-Calvo γc-SCID trial developed leukemia as a result of retroviral insertion of the therapeutic sequence in or near a gene associated with childhood leukemias. Thus a key limitation to the development of human gene therapy remains the lack of safe and efficient methods of gene delivery (Verma and Somia, 1997).

Current gene-delivery methods comprise recombinant viruses, which are used in the majority of clinical trials, and synthetic materials, including peptides, polymers, and liposomes. Although viruses are the most efficient vectors, they often initiate immune responses, are limited in the size of genetic material they can carry, are difficult to produce and purify, and exhibit limited target-cell specificity (or often nonspecificity). Cationic polymers (Felgner and Rolland, 1998; Pack et al., 2005; Smith et al., 1997) have the potential to be nontoxic and nonimmunogenic, are chemically and physically stable, are relatively easy to produce in large quantities, and can be targeted to desired cell types; but in general, they are not efficient enough for clinical use. Even the most efficient polymers are orders of magnitude less efficient than viruses (micrograms of DNA are required to achieve transgene expression comparable to that resulting from a virus suspension containing about 10 picograms of genetic material).

## THE GENE-DELIVERY PROBLEM

To escort genes from a solution (e.g., in a vial) to the cell nucleus, gene-delivery vectors must navigate a series of obstacles, both extracellular and intracellular. Viruses have evolved functions to address each of these challenges, but synthetic vectors generally lack one or several of these functions. These obstacles must all be taken into consideration for the rational design of new materials.

The first set of barriers facing gene-delivery vectors appears in transporting genes from the test tube to the membrane of a target cell. First, the vector must bind and condense plasmid DNA to a sufficiently small size to allow efficient cellular internalization and protect the genes from nuclease degradation. Polycations and DNA spontaneously form tight complexes (polyplexes) through entropically driven electrostatic interactions. The resulting particles typically comprise several DNA molecules and hundreds of polymer chains and range in size from a few tens to several hundred nanometers in diameter. Second, the polyplexes must form a stable solution under physiological conditions, which can often be achieved by coating them with a hydrophilic polymer, such as polyethylene glycol. Third, for

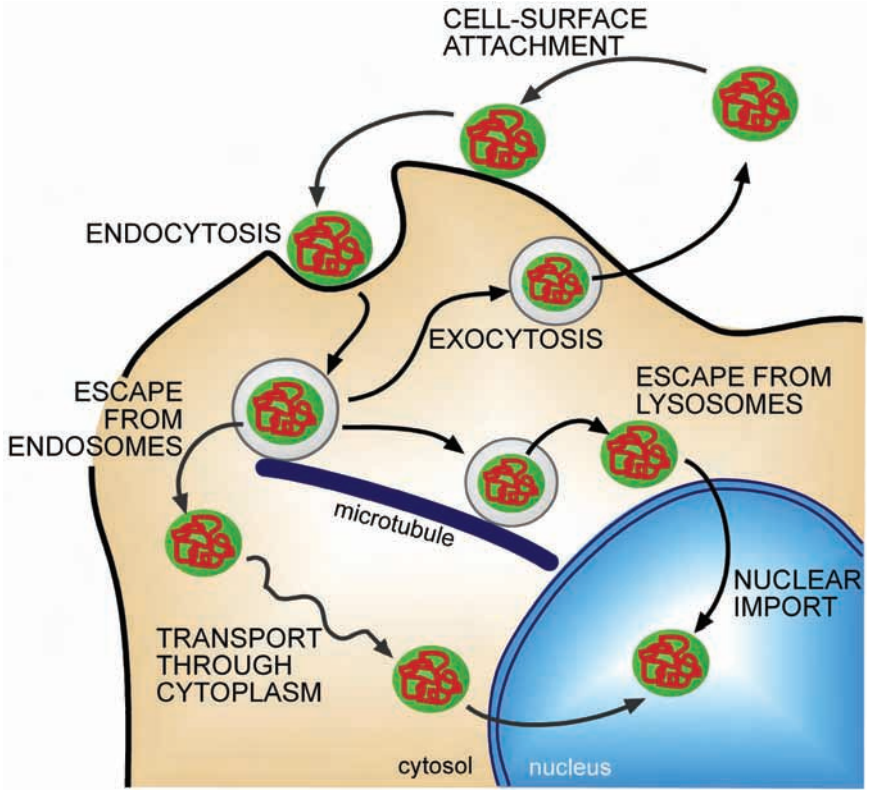


FIGURE 1 Schematic illustration of the main steps in the intracellular processing of polymer-DNA gene-delivery vectors.

many indications, it is critical that vectors recognize specific cells by displaying cell-specific ligands (e.g., small molecules, peptides, proteins, or antibodies).

Following internalization, gene-delivery vehicles must overcome a set of intracellular obstacles (Figure 1), which means the vector must have a functionality to overcome each one. Polyplexes are generally internalized by endocytosis, and once they are in the endocytic pathway, they are routed through a series of vesicles. The typical endpoint of this journey is the lysosome, an acidic vesicle filled with degradative enzymes including nucleases. It is critical, therefore, that DNA and the vector escape these compartments into the cytoplasm.

Next, the vector must escort the DNA through the cytosol toward the nucleus. Because particles as large as typical polyplexes cannot passively diffuse in the cytosol, they require a means of active transport. The genes must then enter the nucleus, with or without the vector material. Although the nuclear envelope con-

tains pores for transporting biomolecules into and out of the nucleus, the process is very tightly regulated to keep out undesirable species, including exogenous genes. Finally, DNA and the vector must separate from one another to allow transcription of the therapeutic gene. The location at which this unpackaging occurs, however, is not generally known. Although more than 95 percent of cells in culture may internalize vectors (on the order of 100,000 copies per cell), typically less than 50 percent express the transgene, suggesting that the majority are lost at one of these steps.

## PROGRESS IN THE DESIGN OF GENE-DELIVERY MATERIALS

Many early studies of gene delivery employed commercially available polymers (Figure 2). Polylysine was one of the first cationic polymers used in the modern era of gene-delivery research (Wu and Wu, 1987; Zauner et al., 1998). Although these studies were promising, it appears unlikely that polylysine-based polyplexes will be clinically useful because of their lack of efficiency.

Polyethylenimine (PEI), however, is one of the most effective gene-delivery polymers (Boussif et al., 1995). Its effectiveness is believed to be due in large part to its efficient escape from the endocytic pathway via the "proton-sponge" mechanism. Because nitrogen represents every third atom in the PEI backbone, this polymer exhibits a very high density of amines, only 15 to 20 percent of which are protonated at physiological pH. As endocytic vesicles are acidified, polyplexes containing PEI (or other proton-sponge materials) are able to buffer the vesicle lumen, leading to an influx of counter ions, osmotic swelling, and vesicle rupture. PEI-mediated gene delivery has been hindered, however, by the polymer's relatively high cytotoxicity in many cell lines, both in culture and in vivo.

In the past two decades, many new types of polymers have been synthesized specifically as gene-delivery vectors. Because polymer-mediated intracellular trafficking is poorly understood, however, many of these designs are based on unproven hypotheses. Results, therefore, have been mixed, with few materials providing highly efficient gene delivery. A current focus in the field, therefore, is developing a new understanding of intracellular processing and polymer structure-activity relationships. Because of space limitations, only a small selection of relevant studies will be described here.

One important approach has been to focus on the synthesis of biocompatible, nontoxic gene-delivery agents, including materials such as poly[α-(4-aminobutyl)-L-glycolic acid] (PAGA), a biodegradable mimic of polylysine (Lim et al., 2000), polyurethanes, disulfide-linked polymers, and poly(β-amino esters) (PBAEs) (Figure 2). As one example of the latter, Forrest et al. (2003) cross-linked low-molecular-weight PEI—which is nontoxic, but ineffective for gene delivery—with small diacrylates (Green et al., 2008). The resulting materials exhibited initial molecular weights sufficient to tightly bind and condense DNA,

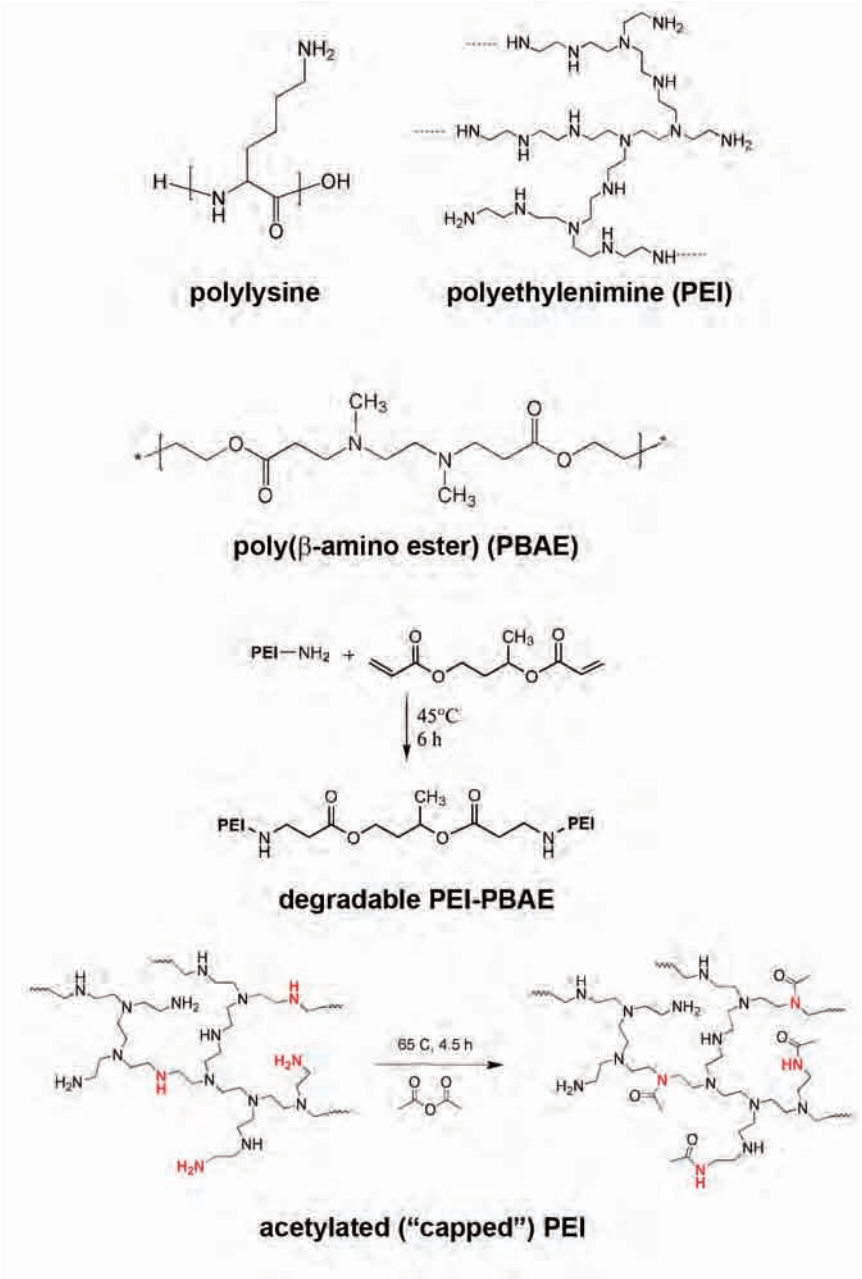


FIGURE 2 Structures of representative gene-delivery polymers.

but they degraded within 8 to 24 hours to nontoxic by-products. These degradable PEI derivatives were as much as 16-fold more efficient than the analogous non-degradable commercial PEI of comparable molecular weight.

PBAEs also have been exploited in combinatorial syntheses in which a panel of diacrylates and amines are cross-linked to generate more than 2,000 unique polymers (Green et al., 2008). These materials have been screened for gene-delivery activity and other important properties, including toxicity. The best polymers were more efficient than the top commercial transfection reagents and, in some situations, were comparable to adenoviruses. Perhaps most important, by correlating gene-delivery activity with polymer/polyplex properties, investigators may begin to extract structure-activity relationships to guide future polymer designs.

Because PEI is an off-the-shelf material, one may also expect that its buffering capacity is not optimal. In fact, Forrest et al. (2004) modified the protonation profile of PEI by reaction with acetic anhydride to convert the primary and secondary amines to secondary and tertiary amides, respectively (Figure 2); such a change should make a poorer proton sponge by decreasing the number of protonable nitrogens in the polymer.

Surprisingly, gene-delivery activity dramatically increased upon acetylation, and the polymer with acetylation on about 57 percent of the primary amines was as much as 60-fold more efficient than unmodified PEI (Gabrielson and Pack, 2006). Subsequent investigation of the mechanisms leading to this unexpected enhancement revealed that PEI acetylation also decreases polymer-DNA binding strength, resulting in enhanced "unpackaging" of polyplexes within target cells. This report was significant in that it identified polymer DNA as a critically important design criterion for gene-delivery materials.

## CONCLUSIONS

A variety of polymers has been used in gene-delivery studies, but they are still orders of magnitude less effective as gene-therapy vectors than viral vectors. As a result, polymers are generally considered unacceptable for clinical applications. Even though the important extra- and intracellular barriers to efficient gene delivery are known, the poor performance of polymer gene-delivery vectors is attributable to a lack of functionality for overcoming at least one of these barriers.

Based on the large number of studies of off-the-shelf gene-delivery polymers, much has been learned about the structure-function relationships of polymer vectors. This knowledge has been applied to the design and synthesis of new polymers, tailor-made for gene delivery, and a number of promising candidates have been reported in recent years. As our understanding of polymer gene-delivery mechanisms improves, it is likely that polymer-based gene-delivery systems will become an important tool in human gene therapy.

## REFERENCES

Blaese, R.M., K.W. Culver, A.D. Miller, C.S. Carter, T. Fleisher, M. Clerici, G. Shearer, L. Chang, Y. Chiang, P. Tolstoshev, J.J. Greenblatt, S.A. Rosenberg, H. Klein, M. Berger, C.A. Mullen, W.J. Ramsey, L. Muul, R.A. Morgan, and W.F. Anderson. 1995. T lymphocyte-directed gene therapy for ADA-SCID: initial trial results after four years. Science 270(5235): 475–480.

Boussif, O., F. Lezoualch, M.A. Zanta, M.D. Mergny, D. Scherman, B. Demeneix, and J.P. Behr. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proceedings of the National Academy of Sciences of the United States of America 92(16): 7297–7301.

Cavazzana-Calvo, M., S. Hacein-Bey, G. de Saint Basile, F. Gross, E. Yvon, P. Nusbaum, F. Selz, C. Hue, S. Certain, J.L. Casanova, P. Bousso, F.L. Deist, and A. Fischer. 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science 288(5466): 669–672.

Felgner, P.L., and A. Rolland. 1998. Non-viral gene delivery systems. Advanced Drug Delivery Reviews 30(1-3): 1–3.

Forrest, M.L., J.T. Koerber, and D.W. Pack. 2003. A degradable polyethylenimine derivative with low toxicity for highly efficient gene delivery. Bioconjugate Chemistry 14(5): 934–940.

Forrest, M.L., G.E. Meister, J.T. Koerber, and D.W. Pack. 2004. Partial acetylation of polyethylenimine enhances in vitro gene delivery. Pharmaceutical Research 21(2): 365–371.

Gabrielson, N.P., and D.W. Pack. 2006. Acetylation of polyethylenimine enhances gene delivery via weakened polymer/DNA interactions. Biomacromolecules 7(8): 2427–2435.

Green, J.J., R. Langer, and D.G. Anderson. 2008. A combinatorial polymer library approach yields insight into nonviral gene delivery. Accounts of Chemical Research 41(6): 749–759.

Kay, M.A., C.S. Manno, M.V. Ragni, P.J. Larson, L.B. Couto, A. McClelland, B. Glader, A.J. Chew, S.J. Tai, R.W. Herzog, V. Arruda, F. Johnson, C. Scallan, E. Skarsgard, A.W. Flake, and K.A. High. 2000. Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. Nature Genetics 24(3): 257–261.

Khuri, F.R., J. Nemunaitis, I. Ganly, J. Arseneau, I.F. Tannock, L. Romel, M. Gore, J. Ironside, R.H. MacDougall, C. Heise, B. Randlev, A.M. Gillenwater, P. Bruso, S.B. Kaye, W.K. Hong, and D.H. Kirn. 2000. A controlled trial of intratumoral ONYX-015, a selectively-replicating adenovirus, in combination with cisplatin and 5-fluorouracil in patients with recurrent head and neck cancer. Nature Medicine 6(8): 879–885.

Lim, Y.-B., C.-H. Kim, K. Kim, S.W. Kim, and J.-S. Park. 2000. Development of a safe gene delivery system using biodegradable polymer, poly[α-(4-aminobutyl)-l-glycolic acid]. Journal of the American Chemical Society 122(27): 6524–6525.

Pack, D.W., A.S. Hoffman, S. Pun, and P.S. Stayton. 2005. Design and development of polymers for gene delivery. Nature Reviews. Drug Discovery 4(7): 581–593.

Smith, J., Y. Zhang, and R. Niven. 1997. Toward development of a non-viral gene therapeutic. Advanced Drug Delivery Reviews 26(2-3): 135–150.

Verma, I.M., and N. Somia. 1997. Gene therapy—promises, problems and prospects. Nature 389(6648): 239–242.

Wu, G.Y., and C.H. Wu. 1987. Receptor-mediated in vitro gene transformation by a soluble DNA carrier system. Journal of Biological Chemistry 262(10): 4429–4432.

Zauner, W., M. Ogris, and E. Wagner. 1998. Polylysine-based transfection systems utilizing receptor-mediated delivery. Advanced Drug Delivery Reviews 30(1-3): 97–113.

# Traceable Drug Delivery: Lighting the Way with Qdots

Xiaohu Gao
*University of Washington*
*Seattle, Washington*

Semiconductor nanocrystals, also known as quantum dots (Qdots), have become an indispensable tool in biomedical research, especially for multiplexed, quantitative, and long-term fluorescence imaging and detection (Alivisatos, 2004; Medintz et al., 2005; Michalet et al., 2005; Smith et al., 2006). The basic rationale for using Qdots is their unique optical properties are generally not available in individual molecules or bulk semiconductor solids. In comparison with conventional organic dyes and fluorescent proteins, Qdots have distinctive characteristics, such as size-tunable light emission, improved signal brightness, resistance against photobleaching, and simultaneous excitation of multiple fluorescent colors.

Recent advances in nanoparticle-surface chemistry have led to the development of polymer-encapsulated probes that are highly fluorescent and stable under complex biological conditions (Dubertret et al., 2002; Gao et al., 2004; Wu et al., 2003). This new generation of water-soluble Qdots has solved the problems of quantum yield decrease, chemical sensitivity, and short shelf-life, which were previously encountered in the ligand exchange-based Qdot solubilization method (Chan and Nie, 1998). As a result, Qdots linked with bioaffinity molecules have created new opportunities for multicolor molecular imaging in living cells and animal models, as well as for traceable drug delivery (Dahan et al., 2003;

Dubertret et al., 2002; Gao et al., 2004; Larson et al., 2003; Lidke et al., 2004; Wu et al., 2003).

## TRACEABLE DRUG DELIVERY

Traceable drug delivery has the potential to elucidate the pharmacokinetics and pharmacodynamics of drug candidates and to provide design principles for engineering drug carriers. Due to concerns about long-term in vivo toxicity and degradation, however, Qdots are currently limited to use in cells and small animals. However, because both cells and small animals are used extensively in testing drug candidates, even in these limited studies, traceable therapeutics have had a significant impact on life-science research, such as in drug discovery, validation, and delivery.

Following drug molecules or drug carriers noninvasively in real time in live organisms requires specialized imaging techniques. Compared with traditional imaging modalities, such as magnetic resonance imaging (MRI) and positron emission tomography (PET), optical imaging is highly sensitive, quantitative, and capable of multiplexing. In addition, it is significantly cheaper than MRI and PET; thus it will substantially reduce the cost and shorten the time of new drug development. Therefore, for the development and optimization of nanocarriers, Qdots can be an excellent "prototype" from which biocompatible carriers of similar sizes and with similar surface properties can be made for clinical uses.

The importance of the structural properties of Qdots for drug-delivery research has only recently been realized. First, the size of Qdots can be continuously tuned from 2–10 nm in diameter, which, after polymer encapsulation, generally increases to 5–20 nm. Particles smaller than 5 nm are quickly cleared by renal filtration (Choi et al., 2007), whereas bigger particles are more likely to be taken up by the reticuloendothelial system before reaching the targeted disease sites. In addition, larger particles have limited penetration into solid tissues. Recent advances in Qdot nanocrystal synthesis will enable scientists to systematically assess the effects of size on delivery efficiency and specificity and identify the optimal dimensions of drug carriers.

Second, because of the high surface-to-volume ratio of nanomaterials, it is possible to link multiple functionalities on single Qdots while keeping the overall size in the optimal range. For example, the Qdot core can serve as the structural scaffold and the imaging contrast agent; and small-molecule hydrophobic drugs can be embedded between the inorganic core and the amphiphilic polymer coating. Hydrophilic therapeutic agents (e.g., small interfering RNA [siRNA] and antisense oligodeoxynucleotide [ODN]) and targeting biomolecules (e.g., antibodies, peptides, and aptamers), in turn, can be immobilized onto the hydrophilic side of the amphiphilic polymer via either covalent or noncovalent bonds. The fully integrated nanostructure may behave like a magic bullet that not only can

identify, bind to, and treat diseased cells, but can also emit detectable signals for real-time monitoring of its trajectory.

## DELIVERY OF siRNA USING QDOTS

RNA interference (RNAi), which is emerging as one of the most powerful technologies for sequence-specific suppression of genes, has potential applications ranging from functional gene analysis to therapeutics. Because of the relatively low immunogenic and oncologic effects of RNAi, the development of nonviral delivery methods in vitro and in organisms is generating considerable interest. In recent years, a number of strategies have been developed based on liposomes, gold and silica nanoparticles, cationic and biodegradable polymers, and peptides (Bielinska et al., 1999; Boussif et al., 1995; Chesnoy and Huang, 2000; Kneuer et al., 2000; Niidome et al., 2004; Roy et al., 2005; Rudolph et al., 2003; Sandhu et al., 2002; Takeshita et al., 2005; Tang et al., 1996; Zanta et al., 1999). The delivery efficiency, however, remains low, especially under in vivo conditions. Another limitation of existing delivery technologies is the lack of an intrinsic signal for long-term, real-time imaging of siRNA transport and release.

We recently developed a new technology by combining Qdots with amphipol, another class of nanomaterial, for traceable and efficient delivery of siRNA molecules. Amphipols, linear polymers with alternating hydrophilic and hydrophobic side chains, are widely used for solubilizing integral membrane proteins and delivering them into cell lipid bilayers (Gorzelle et al., 2002; Nagy et al., 2001; Pocanschi et al., 2006; Tribet et al., 1996, 1997). Unlike detergent-based micelles, amphipols belt around the transmembrane domain of membrane proteins and do not disrupt the integrity of cell membranes during delivery. To our surprise, however, when amphipols are mixed with nanoparticles coated with hydrophobic surface ligands, these two types of nanomaterials form stable complexes that are not only capable of carrying siRNA molecules into cytoplasm but can also protect them from enzymatic degradation.

Compared with classic siRNA carriers, such as Lipofectamine, this new class of nanocarrier works in both serum-free and complete cell-culture media. The new nanocarrier also outperforms polyethyleneimine in gene silencing under both conditions with significantly reduced toxicity. In addition, Qdots provide a bright, stable fluorescent signal for intracellular siRNA imaging (Figure 1).

## CONCLUSION AND PERSPECTIVE

As a powerful imaging probe, Qdots already play an important role in fundamental biology, as well as in in vitro disease diagnostics and prognostics. The unique structural and surface properties of Qdots, such as tunable and uniform size, flexible drug-linking and doping mechanism, large surface-to-volume ratio, and a wide spectrum of surface-reactive groups, have recently opened a new

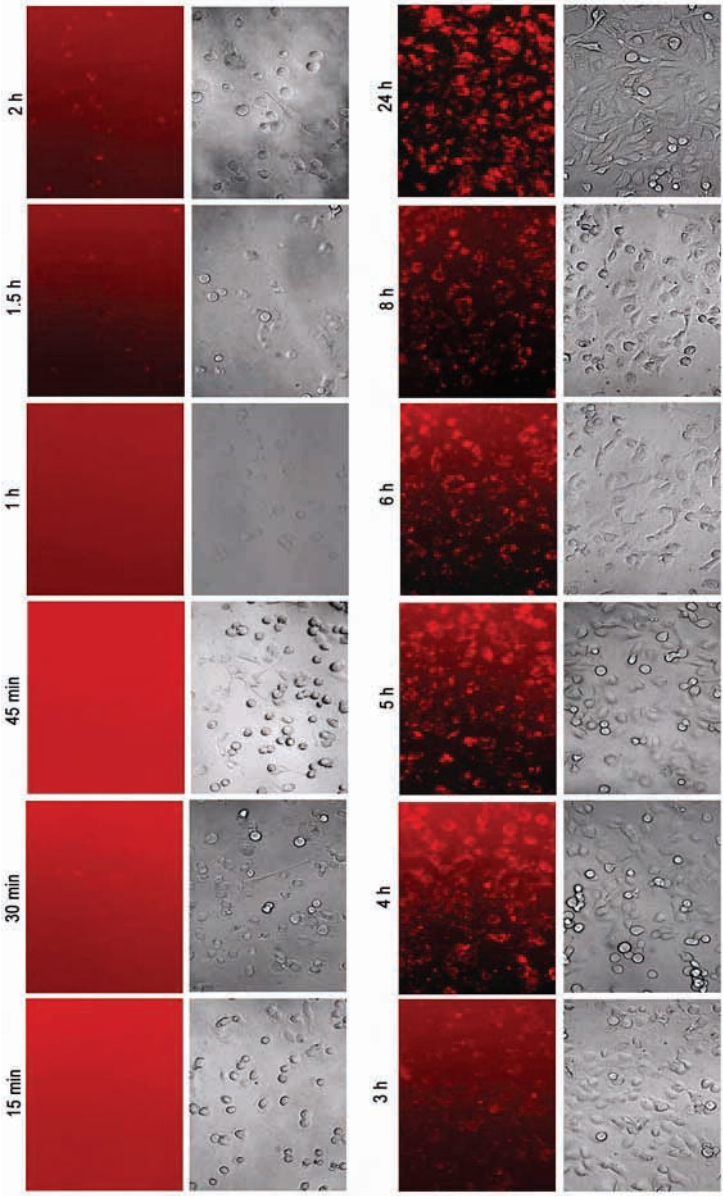


FIGURE 1 Time-dependent fluorescence imaging of the entry and transport of Qdot-siRNA nanoparticles in living cells. Images were obtained 15 minutes to 24 hours after the addition of Qdot-siRNA. Top panels are fluorescence images, and bottom panels are the corresponding bright-field images. (Figure can be viewed in color at *www.nae.edu/frontiers.*)

avenue of research—targeted and traceable drug delivery. However, high-quality Qdots (visible and near infrared dots with a narrow emission profile and high quantum yield) are mainly made with heavy metals whose long-term toxicity is largely unknown. Despite this limitation, they have been used as drug carriers in cells and small animals and have proven to be an outstanding discovery tool for drug screening and validation, as well as a prototype material for drug-carrier engineering. If high-quality Qdots could be prepared from relatively nontoxic compounds (e.g., silicon and carbon), or if the toxic components could be inertly protected from exposure and subsequently cleared from the body, Qdots could become clinically relevant.

Another primary challenge of drug delivery is maintaining a useful concentration of the drug in the targeted tissue while preventing toxicity. Achieving this therapeutic window has not been studied with Qdots thus far, but, ideally, engineered Qdots should be able to stabilize therapeutic compounds, increase their plasma-circulation time while reducing the concentration of free drug to minimize unwanted side effects, and release the drug with a well-controlled profile. In addition, the targeting and therapeutic compounds might be covalently linked to the Qdot surface via cleavable chemical bonds so that the bioconjugates are initially large enough to avoid renal filtration, and later, after the ligands have been cleaved, small enough to be cleared out of the body.

## ACKNOWLEDGMENTS

This work was supported by grants from NIH, NSF, DOD, and the University of Washington. The author thanks NSF for a Faculty Early Career Development Award (CAREER).

## REFERENCES

Alivisatos, P. 2004. The use of nanocrystals in biological detection. Nature Biotechnology 22(1): 47–52.

Bielinska, A.U., C. Chen, J. Johnson, and J.R. Baker, Jr. 1999. DNA complexing with polyamidoamine dendrimers: implications for transfection. Bioconjugate Chemistry 10(5): 843–850.

Boussif, O., F. Lezoualc'h, M.A. Zanta, M.D. Mergny, D. Scherman, B. Demeneix, and J.P.A Behr. 1995. Versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proceedings of the National Academy of Sciences of the United States of America 92(16): 7297–7301.

Chan, W.C.W., and S. Nie. 1998. Quantum dot bioconjugates for ultrasensitive nonisotopic detection. Science 281(5385): 2016–2018.

Chesnoy, S., and L. Huang. 2000. Structure and function of lipid-DNA complexes for gene delivery. Annual Review of Biophysics and Biomolecular Structure 29: 27–47.

Choi, H.S., W. Liu, P. Misra, E. Tanaka, J.P. Zimmer, B. Itty Ipe, M.G. Bawendi, and J.V. Frangioni. 2007. Renal clearance of quantum dots. Nature Biotechnology 25(10): 1165–1170.

Dahan, M., S. Lévi, C. Luccardini, P. Rostaing, B. Riveau, and A. Triller. 2003. Diffusion dynamics of glycine receptors revealed by single-quantum dot tracking. Science 302(5644): 442–445.

Dubertret, B., P. Skourides, D.J. Norris, V. Noireaux, A.H. Brivanlou, and A. Libchaber. 2002. In vivo imaging of quantum dots encapsulated in phospholipid micelles. Science 298(5599): 1759–1762.

Gao, X., Y. Cui, R.M. Levenson, L.W.K. Chung, and S. Nie. 2004. In vivo cancer targeting and imaging with semiconductor quantum dots. Nature Biotechnology 22(8): 969–976.

Gorzelle, B.M., A.K. Hoffman, M.H. Keyes, D.N. Gray, D.G. Ray, and C.R. Sanders. 2002. Amphipols can support the activity of a membrane enzyme. Journal of the American Chemical Society 124(39): 11594–11595.

Kneuer, C., M. Sameti, U. Bakowsky, T. Schiestel, H. Schirra, H. Schmidt, and C.M.A Lehr. 2000. Nonviral DNA delivery system based on surface modified silica-nanoparticles can efficiently transfect cells in vitro. Bioconjugate Chemistry 11(6): 926–932.

Larson, D.R., W.R. Zipfel, R.M. Williams, S.W. Clark, M.P. Bruchez, F.W. Wise, and W.W. Webb. 2003. Water-soluble quantum dots for multiphoton fluorescence imaging in vivo. Science 300(5624): 1434–1436.

Lidke, D.S., P. Nagy, R. Heintzmann, D.J. Arndt-Jovin, J.N. Post, H.E. Grecco, E.A. Jares-Erijman, and T.M. Jovin. 2004. Quantum dot ligands provide new insights into erbB/HER receptor-mediated signal transduction. Nature Biotechnology 22(2): 169–170.

Medintz, I.L., H.T. Uyeda, E.R. Goldman, and H. Mattoussi. 2005. Quantum dot bioconjugates for imaging, labeling and sensing. Nature Materials 4(6): 435–446.

Michalet, X., F.F. Pinaud, L.A. Bentolila, J.M. Tsay, S. Doose, J.J. Li, G. Sundaresan, A.M. Wu, S.S. Gambhir, and S. Weiss. 2005. Quantum dots for live cells, in vivo imaging, and diagnostics. Science 307(5709): 538–544.

Nagy, J.K., A. Kuhn Hoffmann, M.H. Keyes, D.N. Gray, K. Oxenoid, and C.R. Sanders. 2001. Use of amphipathic polymers to deliver a membrane protein to lipid bilayers. FEBS Letters 501(2-3): 115–120.

Niidome, T., K. Nakashima, H. Takahashi, and Y. Niidome. 2004. Preparation of primary amine-modified gold nanoparticles and their transfection ability into cultivated cells. Chemical Communications 7: 1978–1979. DOI 10.1039/b406189f.

Pocanschi, C.L., T. Dahmane, Y. Gohon, F. Rappaport, H.J. Apell, J.H. Kleinschmidt, and J.L. Popot. 2006. Amphipathic polymers: tools to fold integral membrane proteins to their active form. Biochemistry 45(47): 13954–13961.

Roy, I., T.Y. Ohulchanskyy, D.J. Bharali, H.E. Pudavar, R.A. Mistretta, N. Kaur, and P.N. Prasad. 2005. Optical tracking of organically modified silica nanoparticles as DNA carriers: a nonviral, nanomedicine approach for gene delivery. Proceedings of the National Academy of Sciences of the United States of America 102(2): 279–284.

Rudolph, C., C. Plank, J. Lausier, U. Schillinger, R.H. Müller, and J. Rosenecker. 2003. Oligomers of the arginine-rich motif of the HIV-1 TAT protein are capable of transferring plasmid DNA into cells. Journal of Biological Chemistry 278(13): 11411–11418.

Sandhu, K.K., C.M. McIntosh, J.M. Simard, S.W. Smith, and V.M. Rotello. 2002. Gold nanoparticle-mediated transfection of mammalian cells. Bioconjugate Chemistry 13(1): 3–6.

Smith, A.M., S. Dave, S. Nie, L. True, and X. Gao. 2006. Multicolor quantum dots for molecular diagnostics of cancer. Expert Review of Molecular Diagnostics 6(2): 231–244.

Takeshita, F., Y. Minakuchi, S. Nagahara, K. Honma, H. Sasaki, K. Hirai, T. Teratani, N. Namatame, Y. Yamamoto, K. Hanai, T. Kato, A. Sano, and T. Ochiya. 2005. Efficient delivery of small interfering RNA to bone-metastatic tumors by using atelocollagen in vivo. Proceedings of the National Academy of Sciences of the United States of America 102(34): 12177–12182.

Tang, M.X., C.T. Redemann, and F.C. Szoka, Jr. 1996. In vitro gene delivery by degraded polyamidoamine dendrimers. Bioconjugate Chemistry 7(6): 703–714.

Tribet, C., R. Audebert, and J.L. Popot. 1996. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. Proceedings of the National Academy of Sciences of the United States of America 93(26): 15047–15050.

Tribet, C., R. Audebert, and J.L. Popot. 1997. Stabilization of hydrophobic colloidal dispersions in water with amphiphilic polymers: applications to integral membrane proteins. Langmuir 13(21): 5570–5576.

Wu, X., H. Liu, J. Liu, K.N. Haley, J.A. Treadway, J.P. Larson, N. Ge, F. Peale, and M.P. Bruchez. 2003. Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. Nature Biotechnology 21(1): 41–46.

Zanta, M.A., P. Belquise-Valladier, and J.P. Behr. 1999. Gene delivery: a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus. Proceedings of the National Academy of Sciences of the United States of America 96(1): 91–96.

# EMERGING NANOELECTRONIC DEVICES

# Introduction

JIA CHEN
*IBM*
*Yorktown Heights, New York*

VICTOR ZHIRNOV
*SRC*
*Research Triangle Park, North Carolina*

This section focuses on integrating novel nanoscale materials and devices, circuit concepts, and sensor functionalities to develop new technologies for information processing. New concepts for devices, fabrication techniques, and system architectures are emerging from research in nanotechnologies.

Many new ideas have been proposed for post-CMOS technologies, such as molecular electronics, carbon nanotube devices, spin devices, and so on. Most likely, the full potential of these new developments will only be realized in combination with new nanoarchitectures that integrate alternative electronic devices onto a silicon platform. The papers in this section describe work on new emerging nanoelectronic devices and materials.

# The Quest for the Next Information-Processing Technology

JEFFREY J. WELSER
*IBM Almaden Research Center*
*San Jose, California*

In recent semiconductor technology generations, exponentially increasing power density has started to limit the historical benefits of scaling. Thus researchers are looking for entirely new device approaches and methods of computation in emerging nanoscale technologies. The Nanoelectronics Research Initiative (NRI) is taking on the grand challenge of finding a "new switch" that can continue the exponential increase in information-processing capability, which has benefited not only the semiconductor industry, but nearly every aspect of our electronics and information technology-driven modern economy.

## INTRODUCTION

For more than three decades, the semiconductor industry has been driven by its ability to scale the size of the complementary metal oxide semiconductor (CMOS) field-effect transistor (FET), the key building block in modern integrated circuit (IC) chips. This scaling has enabled the industry to pack twice as many FETs onto a chip every 18–24 months, in what has come to be known as "Moore's law" (Moore, 1965). The result has been an exponential increase in the

information-processing capability per unit area on the chip—or more importantly, per dollar.

Recently, exponentially increasing power density, due to both FET leakage currents and switching energy, has limited the continuation of scaling. Recognizing that the fundamental physics of FET operation, rather than fabrication capability, are likely to impose an ultimate limit on continued scaling in the next 10 to 15 years, researchers are now on a quest for new devices that can continue the trends in information-processing performance.

To take on this grand challenge, the NRI (*nri.src.org*) was formed in 2004 as a consortium of Semiconductor Industry Association (SIA) (*www.sia-online.org*) companies to manage a university-based research program as part of the Semiconductor Research Corporation (SRC) (*www.src.org*). Founded by six U.S. semiconductor companies (AMD, Freescale, IBM, Intel, Micron, and TI), NRI partners with the National Science Foundation (NSF), the National Institute of Standards and Technology (NIST), and state governments to sponsor research currently at 35 U.S. universities in 20 states.

## THE NANOELECTRONICS RESEARCH INITIATIVE

The overall goal of NRI is to demonstrate novel computing devices capable of replacing the CMOS FET as a logic switch in the 2020 time frame. To enable the semiconductor industry to continue the historical cost and performance trends for information technology, these new devices must have a significant advantage over ultimate FETs in power, performance, density, and/or cost. To meet these goals, NRI has focused primarily on research on devices that use new computational-state variables beyond electronic charge. In addition, NRI is investigating new interconnect technologies and novel circuits and architectures, including nonequilibrium systems, for exploiting these devices, as well as improved nanoscale thermal management and novel materials and fabrication methods for structures and circuits. Finally, it is hoped that these technologies will be capable of integrating with CMOS so that their potentially complementary functionality can be exploited in heterogeneous systems and can enable a smooth transition to a new scaling path.

## PHYSICS OF A LOGIC SWITCH

As outlined in the 1970s (Dennard et al., 1974), if one shrinks the critical dimensions of an FET by a factor kappa, while simultaneously increasing the doping levels and decreasing the applied voltages by the same factor, the scaled transistor switches faster but consumes half the power and takes up half the area. This means that twice as many transistors can fit in the same area without decreasing power density—the primary reason scaling works without melting the chip.

Leakage currents, however, increase as dimensions shrink; leakage occurs along the transistor channel when the switch is turned off and across the gate insulator, which has become so thin (<1.5 nm, or just a few atomic layers) that quantum tunneling dominates. The leakage power is now becoming equivalent to the active switching power of the transistor.

The 2001 International Technology Roadmap for Semiconductors (ITRS) Emerging Research Device Technical Working Group conducted a highly simplified analysis of a generic electronic switch at thermal equilibrium (Zhirnov et al., 2003). The switch was modeled as a potential barrier separating two quantum wells, corresponding to the simplest version of an FET channel between source and drain contacts. The analysis showed that the channel could conceivably be scaled down to ~1.5 nm and that the transistor could have a minimum switching speed of ~40 fs—significantly smaller and faster than today's FETs of about 30 nm in channel length with ~1 ps switching time. However, to avoid leakage over the barrier at room temperature, the voltage could not be scaled as rapidly as the physical dimensions, and the resulting power density for these switches at maximum packing density would be on the order of 1 MW/cm$^2$—orders of magnitude higher than the practical air-cooling limit of ~100 W/cm$^2$.

Because the theory does not take into account the materials or structures used, it is applicable to any switch that moves charge dissipatively across a potential barrier. This leads to two implications: (1) Simply shrinking an FET to the far nanoscale will not necessarily continue to give the historical benefit of scaling, because the increasing power density will require trading off switching speed for packing density. (2) The existing Si FET road map is likely to reach the minimum practical dimensions in the next 10 to 15 years (ITRS, 2007); although new FET materials or geometries can improve performance, they will not alter the ultimate scaling limits.

## NANOELECTRONICS RESEARCH AREAS

We need a new "switch" for information processing to significantly extend the scaling path. To define research directions, groups from industry, government, and academia have participated in workshops sponsored by SRC, NSF, and SIA (Cavin and Zhirnov, 2004; Cavin et al., 2005, 2006). Thirteen research vectors were defined, and the top five comprise the NRI research program.

### New Devices: Alternative Computational State Vectors

An FET device moves electrons dissipatively to charge (discharge) capacitors to represent a binary "1" ("0"). NRI research focuses on finding alternative ways to represent these states or information in general. Any physical property that can be placed into two or more distinguishable states could potentially be used to represent information. One example is to use the spin of an electron, with spin "up"

representing "1" and spin "down" representing "0." Spin is already used successfully in memory and storage devices, so the challenge is injecting, manipulating, and reading out the spin state of an electron or collection of electrons to build logic gates and circuits. Many different devices are being considered under the broad heading of "spintronics" (Zutic et al., 2004).

Many other materials offer different potential states that could be exploited for logic, including ferroelectric, antiferroelectric, ferromagnetic, antiferromagnetic, ferrotoroidic, ferroelastic, and ferrimagnetic materials (Eerenstein et al., 2006). Even the physical movement of atoms could be considered as a new state variable. Even though atoms are more massive than electrons, it would only be necessary to move them on the order of angstroms to cause large changes in a material (e.g., changing the dipole in a ferroelectric or changing barrier heights at an interface), so that the speed and energy could still be reasonable. And for nanoscale devices, more massive particles are less likely to lose their state by tunneling (Zhirnov and Cavin, 2008).

Much of the work on new state variables relies on the development of new materials. Dilute magnetic semiconductors have the potential to introduce spin into semiconductors (Pearton et al., 2004). Multiferroic materials, which could couple ferroelectric and ferromagnetic parameters, could be used to manipulate spins without magnetic fields (Eerenstein et al., 2006).

A recently discovered material of particular interest is graphene—a single monolayer of graphite with unique transport properties (Geim and Novoselov, 2007). The two-dimensional honeycomb lattice of graphene gives rise to a conical band structure that leads to electrons behaving as massless Dirac fermions. Graphene could not only improve FET devices due to its high carrier velocity, but could also enable new devices exploiting its unique physics. The pseudospin property (Min et al., 2008), for example, could potentially enable a correlated shift of charge density between two graphene layers, which could lead to a new low-energy switch.

## New Methods of Computation: Nonequilibrium Systems

Operating an FET at room temperature requires energy barriers of sufficient height to maintain distinguishability between states, which will also be a factor for any other device. One way to get around this problem would be to recapture the computation energy, rather than allowing it to dissipate as heat. This is the goal of adiabatic or "reversible" computation (Bennett, 1988). Another way would be to do out-of-equilibrium computation. In the solid state, local distributions of carriers (or spins or other phase states) can be out of equilibrium with the ambient "temperature" for a period of time before relaxing to the lattice temperature through phonon collisions or other coupling parameters. If the relaxation time is sufficiently long, the potential barriers could be lower, allowing state manipulation with lower switching energy.

It has not yet been proven experimentally that either approach can be used for computation with reduced energy dissipation, but out-of-equilibrium behavior is a primary motivation for considering alternative-state variables for information processing in the first place. If an alternative-state variable obeys the same Boltzman statistics as a dissipative electron-based system, it will have little chance of offering substantial energy advantages over FETs.

### New Ways of Connecting Devices: Noncharge Data Transfer

Any computation system requires connecting multiple devices and transferring information between them. Charge is the natural carrier to use in electron devices, but this is what drives much of the power consumption in modern ICs. An alternative device should transmit the new state variable since converting back to charge would negate any advantage of the new information token.

For example, a spin device should transmit spin to the next device, such as through a spin-wave bus (Khitun et al., 2007). The movement of a spin wave can be very low energy (Bernevig, 2006), but if electrons must be moved to move the spin, the power advantage is lost. Similarly, if a device uses ferromagnetic or ferroelectric orientation, that should be transmitted through some low-energy magnetic or lattice interaction. Transmitting information short distances by coherent waves or collective effects is a promising idea for interconnecting devices, which favors architectures based on nearest-neighbor device coupling. However, it will still probably be necessary to convert to charge for cross-chip interconnects and for coupling back out to the external world.

### New Methods of Managing Heat: Nanoscale Phonon Engineering

Finding more energy-efficient ways to cool devices, given the immediate importance for current CMOS chips, is a very active area of research in the semiconductor industry. The NRI focus is limited to looking at ways of controlling phonon flow for more efficient phonon extraction and manipulation in device structures. This is being coupled with work on the nonequilibrium system, because finding ways to lengthen the time the state is out of equilibrium with the thermal environment could be key to the development of low-energy computation.

It may even be possible to use phonons themselves as the state variable (Wang and Li, 2007). Given the large costs in energy required to lower temperatures, current research is focusing on room-temperature operation. If exceptionally efficient cooling mechanisms were discovered, however, that boundary condition could change.

### New Fabrication Methods: Directed Self-Assembly of Devices

Directed self-assembly combines traditional patterning with self-organizing systems to create nanostructures. The ability to fabricate nanoscale CMOS cost-effectively is a challenge for the industry, and the goal of focusing on self-assembly is to improve lithography for continued scaling. NRI work is focused only on self-assembly for the direct creation of new device structures, such as arrays of self-assembled magnetic dots for magnetic quantum cellular automata circuits (Bernstein et al., 2005; Liu and Reinke, 2008). A wider focus on fabrication is likely in the future, once the state variable and architecture have been established.

## SUMMARY

The daunting grand challenge of finding a device capable of extending information processing beyond the ultimate limits of CMOS technology is similar to the challenge faced in the 1940s when solid-state transistors were developed to replace vacuum tubes. The current NRI program (Welser et al., 2008) is largely focused on the first research vector—finding a new device—which would more clearly define directions for research on the other vectors. However, the investigation of alternative devices, data transport, thermal transport, and manufacturability will ultimately have to be tightly integrated to bridge the gap from basic science to a practicable information-processing technology.

The CMOS FET is a very efficient switch, and the limits it will approach or reach in the next decade are fundamental to any device operating at room temperature. Although it is difficult to predict what device might be capable of surmounting these limits, a few educated guesses can be made. To reduce power, the new device will probably be slower and will rely on local interconnects. To compensate, it will have to be densely packed and will probably have a three-dimensional architecture. Finally, cost-effective manufacturing will favor uniform arrays of devices—potentially self-assembled—that are robust at the device or architectural level to the increasing variability at the nanoscale.

What information-processing architecture is capable of using such a device? The brain, of course, is often cited as a proof of concept that such a device can be found, at least for certain types of applications. The brain is extremely efficient at pattern recognition, for example, but not particularly good at the mathematical computation used in most of our electronic systems today. Hence the quest for a new information-processing technology will require not only finding a new device, but also rethinking how to apply that device and architecture to new applications and products in the future.

## ACKNOWLEDGMENTS

The author gratefully acknowledges the work of the many contributors from industry, academia, NSF, NIST, and other government agencies who worked on this topic in the SIA Technical Subcommittee, the Silicon Nanoelectronics and Beyond working groups, and the ITRS Emerging Research Devices and Emerging Research Materials subgroups. In particular, the author would like to thank G. Bourianoff of Intel and R. Cavin and V. Zhirnov of SRC for their contributions.

## REFERENCES

Bennett, C.H. 1988. Notes on the history of reversible computation. IBM Journal of Research and Development 32(1): 16–23.

Bernevig, B.A., and S. Zhang. 2006. Toward dissipationless spin transport in semiconductors. IBM Journal of Research and Development 50(1): 141–148.

Bernstein, G.H., A. Imre, V. Metlushko, A. Orlov, L. Zhou, L. Ji, G. Csaba, and W. Porod. 2005. Magnetic QCA systems. Microelectronics Journal 36(7): 619–624.

Cavin, R.K., and V.V. Zhirnov. 2004. Silicon nanoelectronics and beyond: reflections from a semiconductor industry-government workshop. Journal of Nanoparticle Research 6(2): 137–147.

Cavin, R.K., V.V. Zhirnov, G.I. Bourianoff, J.A. Hutchby, D.J.C. Herr, H.H. Hosack, W.H. Joyner, and T.A. Wooldridge. 2005. A long-term view of research targets in nanoelectronics. Journal of Nanoparticle Research 7(6): 573–586.

Cavin, R.K., V.V. Zhirnov, D.J.C. Herr, A. Avila, and J. Hutchby. 2006. Research directions and challenges in nanoelectronics. Journal of Nanoparticle Research 8(6): 841–858.

Dennard, R.H., F.H. Gaensslen, H.-N. Yu, V.L. Rideout, E. Bassous, and A.R. LeBlanc. 1974. Design for ion-implanted MOSFET's with very small physical dimensions. IEEE Journal of Solid-State Circuits SC-9(5): 256–268.

Eerenstein, W., N.D. Mathur, and J.F. Scott. 2006. Multiferroic and magnetoelectric materials. Nature 442(7104): 759–765.

Geim, A.K., and K.S. Novoselov. 2007. The rise of graphene. Nature Materials 6(3): 183–191.

ITRS (International Technology Roadmap for Semiconductors). 2007. ITRS 2007 Edition. Available online at <www.itrs.net/Links/2007ITRS/Home2007.htm>.

Khitun, A., D.E. Nikonov, M. Bao, K. Galatsis, and K.L. Wang. 2007. Efficiency of spin-wave bus for information transmission. IEEE Transactions on Electron Devices 54(12): 3418–3421.

Liu, H., and P. Reinke. 2008. Formation of manganese nanostructures on the Si(100)-(2x1) surface. Surface Science 602(4): 986–992.

Min, H., G. Borghi, M. Polini, and A.H. MacDonald. 2008. Pseudospin magnetism in graphene. Physical Review B 77(041407):1–4.

Moore, G.E. 1965. Cramming more components onto integrated circuits. Electronics 38(8): 114–117.

Pearton, S.J., W.H. Heo, M. Ivill, D.P. Norton, and T. Steiner. 2004. Dilute magnetic semiconducting oxides. Semiconductor Science and Technology 19(10): R59–R74.

Wang, L., and B. Li. 2007. Thermal logic gates: computation with phonons. Physical Review Letters 99(177208): 1–4.

Welser, J.J., G.I. Bourianoff, V.V. Zhirnov, and R.K. Cavin. 2008. The quest for the next information processing technology. Journal of Nanoparticle Research 10(1): 1–10.

Zhirnov, V.V., and R.K. Cavin. 2008. Charge of the heavy brigade. Nature Nanotechnology 3(7): 377–378.

Zhirnov, V.V., R.K. Cavin, J.A. Hutchby, and G.I. Bourianoff. 2003. Limits to binary logic switch scaling—a gedanken model. Proceedings of the IEEE 91(11): 1934–1939.

Zutic, I., J. Fabian, and S. Das Sarma. 2004. Spintronics: fundamentals and applications. Reviews of Modern Physics 76(2): 323–410.

# Molecular and Polymer Nanodevices

Nikolai Zhitenev
*National Institute of Standards and Technology*
*Gaithersburg, Maryland*

The research and engineering community has been looking intensively for possibilities of extending information-processing technologies into the post-CMOS era. Recently, the Nanotechnology Research Initiative (NRI), a consortium of leading semiconductor companies, has formulated a set of research priorities (Welser et al., 2008) based on an analysis of the ultimate limitations of the present technology and trends in research and development. One of the recommended vectors is devices that operate with state variables different from an electronic charge. One possibility would be a solid-state switch with the computational state defined by the spatial locations of heavy particles, such as ions, atoms, or molecular conformations.

The potential advantage of a heavier information carrier can be easily illustrated (Cavin et al., 2006). The scaling of CMOS devices operating with electronic charge will eventually reach their limit when the logic or memory state decays because of electron tunneling under the barriers. For a given barrier height and width limited by material constraints and device size and by the requirement of minimal power dissipation, carriers such as ions or atoms, which are thousands of times heavier than electrons, can provide much greater stability to the compu-

tational state. Ironically, the use of heavy carriers is impractical in larger devices because of their limited mobility. However, in a nanometer-sized device, the ion/atom transport can be fast enough for practical applications.

Short molecules and macromolecules can be used as active materials for heavier switching devices. Devices built with short molecules have long been considered promising candidates for the post-CMOS era for a number of reasons. First, organic molecules can be extremely small and at the same time exactly reproducible as stand-alone units. In addition, numerous synthetic techniques have been developed, and the variety of organic compounds is enormous. Some well-known approaches to molecular electronics already rely on molecular conformations or oxidation to achieve electronic functionality (e.g., Chen et al., 1999; Collier at al., 2000).

However, the reliable fabrication of devices and the assembly of molecules into circuits turn out to be extremely challenging. In this presentation, I will describe some examples from our research that illustrate some of the challenges of fabricating and characterizing molecular devices. Before we began designing a molecular switch or transistor, we tested simpler building blocks in the molecular "tool box," such as molecular "wires" and molecular "barriers."

The investigation of the electronic properties of molecular devices is intimately related to research on alternative fabrication routes that can be compatible with the new materials. First, the required feature size is often beyond the limits of the best lithography machines. Second, the properties of pristine material can be substantially altered by, for example, exposure to a high-energy electron beam encountered in the e-beam lithography step, etching, or contact deposition.

## NEW FABRICATION TECHNIQUES

In our research, we focused on the noninvasive fabrication of nano- and mesoscale molecular devices and the effects of fabrication on their structural and electronic properties (Zhitenev et al., 2006). We fabricated metal-molecular monolayer-metal junctions using three complementary original techniques that target different fabrication issues. After screening many possible candidates for molecular "wires" and molecular "barriers," we selected representative molecules capable of forming a dense self-assembled monolayer (SAM) with the most robust structural and electrical properties.

The first technique targeted nearly single-molecule devices. The junctions were formed on the surface of the tips to exploit the evaporation of contacts from different angles with an assembly of SAMs in the middle. Device conductance was monitored during the formation of the junction. Devices were studied at multiple stages, from minimally detectable conductance below the conductance level of a single-molecule junction to an approximately single-molecule device to a multimolecule device. The shortcoming of this technique was that it relied

on conductance as a single feedback parameter for characterizing the junction formation.

The second technique was a planar evolution of the first one. In this case, we used nanoscale stencil masks prefabricated on chip surfaces and angled evaporations to fabricate the molecular junctions with well-defined areas ranging from 30 nm × 30 nm to 1 μm × 2 μm.

The third technique, nanotransfer, was designed to avoid the evaporation on top of the SAM and hence to examine the potential damage of evaporation. A column pattern was fabricated on a flexible polymer (PDMS) stamp, and thin metal (gold) was evaporated on the stamp. Functional groups at the top surface of the SAM bonded to the gold when the stamp was brought into contact with the molecular layer. Gold dots with diameters of 20 to 100 nm transferred to the top of the SAM served as contacts to the molecular junction, which was probed by a conducting atomic-force microscope.

## DISCUSSION

Using these three techniques, we were able to examine a variety of fabrication issues, materials properties, and transformations affecting the apparent electrical behavior of molecular devices. We found that it was extremely difficult to fabricate junctions without defects, which can arise from a variety of origins and have a range of effects on device performance. In general, metal electrodes have surface topographical features comparable to or longer than the molecules. Thus the order of molecular assembly was disrupted at the grain boundaries making this location "defective" in a structural and electronic sense.

In addition, the nucleation of metal films from the evaporation stream, the resulting surface morphology, the penetration of metal atoms and particles into the molecular layer, and chemical reactions of metals and molecules with oxygen and water were specific to the particular combination of metal and molecular species. For example, junctions with molecular "wires" can appear more resistive than junctions with molecular "barriers" because of the different penetration of gold clusters from the top electrode into the film. If the top electrode is titanium or nickel or another reactive metal, all distinctions between "wires" and "insulators" are lost, because the entire molecular layer is converted into metal carbides or oxycarbides.

In general, the electronic levels of molecules were strongly shifted from the Fermi level of metal electrodes, typically by 1 to 5 eV, and the tunneling conductance of such mismatched systems was too low for practical applications. Defects that created electronic states 50 to 200 meV from the Fermi energy level contributed significantly to electronic transport and defined overall behavior. The reliance on precise atomic positions of the device constituents generally failed because the defects took over.

## NEXT STEPS

The question is whether we really need atomic precision to build functional devices. The latest research and engineering results (Scott and Bozano, 2007; Waser and Aono, 2007) have shown that when the overall properties are defined by the self-averaging of multiple imprecise events, the traditional "statistical" approach can be extended to very small devices.

For example, we found useful switching functionality in polymer films and the monolayers of macromolecules (Zhitenev et al., 2007) that were just a little bit thicker (5–12 nm) than molecular monolayers (1–2.5 nm). The resistive switching was the result of the shift of electronic levels in the device caused by ionic motion. There is nothing precise about a single-ion position or motion, but the total number of ions in the nanoscale device was large enough to result in statistically reproducible switching.

Initially, the devices were nonconducting, but when applied voltage exceeded some threshold level, the devices switched to a conductive state, which was stable at small applied voltage levels. Higher voltage levels of the opposite polarity switched the devices back to a nonconductive state. The switching voltage scaled linearly with the film thickness and depended on the concentration of ionic groups in the film. In addition, ionic groups could be modified in a straightforward way by partly or fully replacing protons with other ions.

Because many of these substitutions were reversible, we were able to examine multiple chemical compositions with the same mesoscopic devices. When we did so, we found that the chemical modifications had major effects on switching behavior. A simple physical model that captures the most essential experimental finding is described below.

In the "off" state, all molecular energy levels are a few electrovolts from the electrode Fermi level, and switching the electric field is strong enough to break the ionic bonds. The ion separation acts as internal "chemical gating," shifting some energy levels into better alignment with the Fermi energy of the electrodes. These electronic states, with energy levels of 100–300 meV from the Fermi level, form the conducting channels. A strong electric field of the opposite polarity pushes the separated cations back, facilitating the recombination with anions at the polymer backbone and eliminating the conductive electronic levels. The electric field required for ion separation depends on the size and properties of the ions. For example, larger monovalent cations have smaller bond strength. Thus the devices can be turned on at smaller threshold voltage. With multivalent ions forming chemical bonds to two or more anion groups, the threshold voltage is significantly increased.

Polymer switches are just one example of material systems that display resistive bistability. There are many other candidates based on various organic and inorganic compositions (Scott and Bozano, 2007; Waser and Aono, 2007). For all of these materials, there is a common element in switching behavior. The

switching functionality is caused by the movement or displacement of heavy particles, such as ions or atoms, over distances ranging from an elementary cell to the size of an entire device. The variety of materials capable of displaying switching leads us to believe that devices based on atom/ion motion can eventually be used in practical circuits.

## CONCLUSION

The use of switching in memory and storage devices has been the main driver for the development of switches by most of the major semiconductor companies. However, if the switching phenomenon can be reliably engineered in devices of sufficiently small size, this will lead to the emergence of new hybrid logic circuits based on novel architectural concepts (Strukov and Likharev, 2007). Some of these concepts mimic the "architecture" and the well-developed connectivity of the human brain, an integral combination of memory, connectivity, and computational elements.

## REFERENCES

Cavin, R.K., V.V. Zhirnov, D.J.C. Herr, A. Avila, and J. Hutchby. 2006. Research directions and challenges in nanoelectronics. Journal of Nanoparticle Research 8(6): 841–858.

Chen, J., M.A. Reed, A.M. Rawlett, and J.M. Tour. 1999. Large on-off ratios and negative differential resistance in a molecular electronic device. Science 286(5444): 1550–1552.

Collier, C.P., G. Mattersteig, E.W. Wong, Y. Luo, K. Beverly, J. Sampaio, F.M. Raymo, J.F. Stoddart, and J.R. Heath. 2000. A [2]catenane-based solid state electronically reconfigurable switch. Science 289(5482): 1172–1175.

Scott, J.C., and L.D. Bozano. 2007. Nonvolatile memory elements based on organic materials. Advanced Materials 19(11): 1452–1463.

Strukov, D.B., and K.K. Likharev. 2007. Defect-tolerant architectures for nanoelectronic crossbar memories. Journal of Nanoscience and Nanotechnology 7(1): 151–167.

Waser, R., and M. Aono. 2007. Nanoionics-based resistive switching memories. Nature Materials 6(11): 833–840.

Welser, J.J., G.I. Bourianoff, V.V. Zhirnov, and R.K. Cavin. 2008. The quest for the next information processing technology. Journal of Nanoparticle Research 10(1): 1–10.

Zhitenev, N.B., W.R. Jiang, A. Erbe, Z. Bao, E. Garfunkel, D.M. Tennant, and R.A. Cirelli. 2006. Control of topography, stress and diffusion at molecule-metal interfaces. Nanotechnology 17(5): 1272.

Zhitenev, N.B., A. Sidorenko, D.M. Tennant, and R.A. Cirelli. 2007. Chemical modification of the electronic conducting states in polymer nanodevices. Nature Nanotechnology 2(4): 237–242.

# Roll Printing of Crystalline Nanowires for Integrated Electronic and Sensor Arrays

Ali Javey
*University of California at Berkeley and Lawrence Berkeley National Laboratory*

Zhiyong Fan
*University of California at Berkeley*

Johnny C. Ho
*University of California at Berkeley*

Roie Yerushalmi
*Hebrew University of Jerusalem*

Fabrication of printable sensor arrays on bendable/flexible substrates may enable the development of a wide range of new technologies, including flexible displays, radio frequency identification tags, sensor tapes, artificial skin, and more (Friedman et al., 2005; Huang et al., 2001; Lee et al., 2005; McAlpine et al., 2005; Reuss et al., 2005; Service, 2000; Someya and Sakurai, 2003). Tremendous progress has been made in this field in the past decade, mainly through the exploration of organic materials as active semiconductor components. However, the short lifetimes and low carrier mobility of these materials, as compared to crystalline inorganic semiconductors, have been major obstacles to applications that require high speed, low power, and long lasting electronics (Reuss et al., 2005; Service, 2000; Someya and Sakurai, 2003). Therefore, a new printable electronic materials technology with improved performance and air stability is of great interest for the future of printable electronics.

Recently, new methods of "printing" microscale and nanoscale inorganic structures have been proposed and developed. Unlike their organic counterparts, inorganic materials provide air stability as well as high performance (Ahn et al., 2006; Bryllert et al., 2006; Fan et al., 2008a,b; Ford et al., 2008; Friedman et al.,

2005; Huang et al., 2001; Javey et al., 2007; Lee et al., 2005; McAlpine et al., 2005; Wang et al., 2007; Yerushalmi et al., 2007). One such inorganic material is the crystalline semiconductor nanowire (NW). In this paper, we review recent advances in the assembly and integration of NW arrays on foreign substrates that can be integrated into electronic devices and sensors.

## CRYSTALLINE SEMICONDUCTOR NANOWIRES AS BUILDING BLOCKS FOR ELECTRONICS AND SENSORS

To date, a variety of functional NWs have been synthesized and integrated as building blocks of single-component devices, such as field-effect transistors (FETs), sensors, photodiodes, and electromechanical systems, to mention just a few (Ahn et al., 2006; Bryllert et al., 2006; Fan et al., 2008a,b; Ford et al., 2008; Friedman et al., 2005; Huang et al., 2001; Javey et al., 2007; Lee et al., 2005; McAlpine et al., 2005; Wang et al., 2007; Yerushalmi et al., 2007). These chemically derived single-crystalline nanostructures (the majority of them synthesized by chemical vapor deposition [CVD]) have unique advantages over conventional semiconductors. They enable the integration of high-performance device elements on virtually any substrate (including mechanically flexible plastics) with scaled on-currents and switching speeds comparable to or higher than those of state-of-the-art, planar silicon (Si) structures.

For example, *p*-type FETs based on heterostructured Ge/Si NWs and n-type FETs based on InAs NWs have demonstrated a carrier mobility about 10 times higher than that of Si transistors (Bryllert et al., 2006; Ford et al., 2008; Xiang et al. 2006). These high-mobility NW materials are ideal platforms for high-performance, printable electronics. Uniquely, the electrical properties of NWs are extremely sensitive to their chemical/biological and electromagnetic surroundings because of their miniaturized dimensions, large surface-area-to-volume ratio, and finite carrier concentration. As a result, sensors based on NWs are also highly sensitive. For example, NWs made of Si and $In_2O_3$ have been extensively studied for use in biological and chemical sensors capable of detecting analytes down to the level of single molecules (Zhang et al., 2004; Zheng et al., 2005). CdSe and ZnO NWs, which are optically active and have been investigated in the past, have demonstrated a significantly higher photoresponse than their thin-film or bulk counterparts (Fan et al., 2008a; Yu et al., 2008).

Although NWs are obviously promising materials for high-performance nanoelectronics and sensors, a major challenge to their integration into large-scale devices/circuits is perfecting their controlled assembly on substrates. In recent years, many approaches have been investigated with varying degrees of success. These approaches include liquid-flow alignment, Langmuir-Blodgett technique, alternating current (AC) dielectrophoresis, blown-bubble method, contact and roller printing, and others. In this article, we review recent progress on a highly

efficient, scalable approach for the ordered, uniform assembly of NW arrays on substrates for integration in multifunctional circuits.

## ROLL PRINTING OF NANOWIRES ON SUBSTRATES

We recently developed an NW roll-printing technology to address the need for large-scale assembly of aligned NW arrays on foreign substrates (Fan et al., 2008b; Yerushalmi et al., 2007). The overall process involves (1) optimized catalytic growth of the desired crystalline NWs by CVD on a cylindrical substrate (i.e., roller), and (2) patterned transfer of NWs directly from the roller to a receiver substrate via differential roll printing, as illustrated in Figure 1.

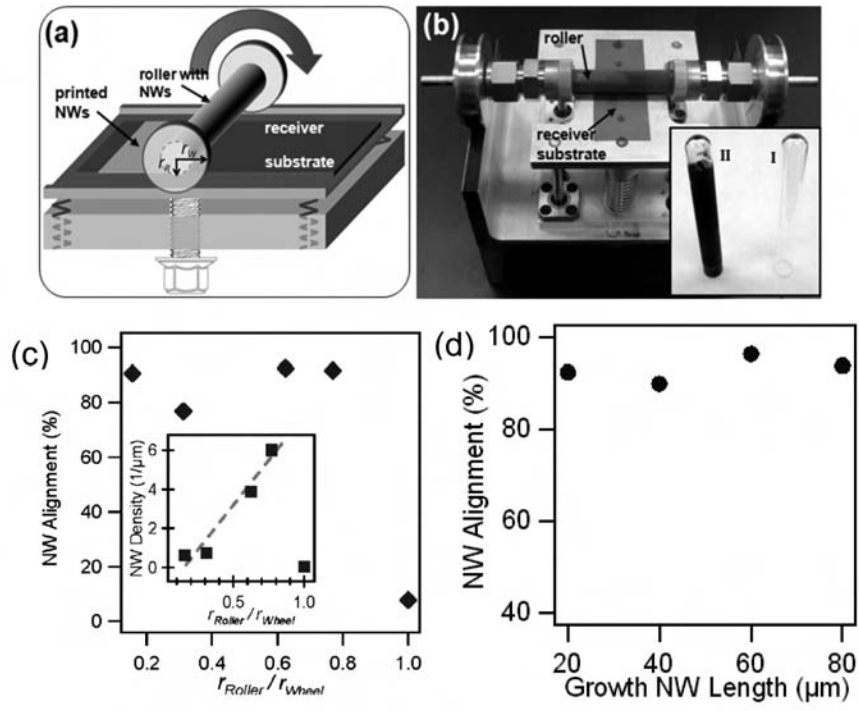


FIGURE 1 Differential roll printing of NWs. (a) Schematic drawing of the printing setup. (b) Optical photograph of the assembled apparatus (top view). The inset shows the blank and NW-coated glass tubes used as rollers (I and II, respectively). (c) The NW alignment and density (inset) as a function of roller-to-wheel size ratio. (d) The alignment of the printed film is nearly independent of NW length. Source: Yerushalmi et al., 2007. Reprinted with permission.

The grown NWs stick out of the surface of the roller with random orientation. The length of the NWs is controlled by the growth time and is typically 20–80 μm for optimal printing results; the diameter (10–100 nm) is controlled by the size of the catalytic nanoparticles used as seeds for CVD growth. The roller is connected to a pair of rotating wheels and brought into contact with a stationary receiver substrate. As the roller is turned under a constant pressure and at a constant speed, NWs are transferred to the receiver substrate, which is coated with a photolithographically patterned photoresist layer that enables the patterned assembly of NWs (Yerushalmi et al., 2007).

An important aspect of this printing process is the mismatch between the radius of the roller and the radius of the wheel ($r_R$, $r_W$, respectively), which causes a shear motion of the roller on the stationary substrate in addition to the rolling motion (Yerushalmi et al., 2007). In traditional roll-printing methods, such a mismatch would be highly undesirable and would distort the printed features. However, the relative sliding motion caused by the mismatch generates the required directing field and shear force to effectively "comb" the NWs, resulting in aligned transfer to the receiver substrate. Without the shear force, a negligible number of NWs are transferred, and their alignment is random, as shown in Figure 1c. This is consistent with the hypothesis that, as randomly aligned NWs on the growth substrates are dragged across the surface of the receiver substrate, they become aligned by mechanical combing.

Once the NWs are anchored by van der Waals forces, they are detached from the growth substrate and transferred to the receiver substrate. Interestingly, the density of the printed NWs shows a near linear dependence on $r_R/r_W$ for $r_R/r_W<1$, as shown in the inset of Figure 1c. This trend is to be expected because the total number of NWs available for transfer is $(2\pi r_R)nW$, where $n$ is the density of NWs on the roller substrate and $W$ is the width of the contact area. Since the printed area covered per revolution is $(2\pi r_W)W$, the maximum printed density is $n(r_R/r_W)$. If we compare the slope of the density of printed NWs with $r_R/r_W$, we get $n$~9 NW/μm (Yerushalmi et al., 2007).

We have observed that, in the range of 20–80 μm, the length of as-grown NWs does not change the printing alignment significantly, as shown in Figure 1d. The high degree of alignment (~90 percent) is independent of the length of the NW and is highly favorable for the scalability of device applications (Fan et al., 2008b). During the printing process, NWs are assembled on both the photoresist and patterned regions of the substrates. The patterned photoresist is later removed by a standard liftoff process using a solvent, leaving behind assembled NWs at the predefined locations (Fan et al., 2008b; Yerushalmi et al., 2007).

This process can be used for a wide range of NW materials, including Si, Ge, and compound semiconductors, and for the entire NW diameter range (10–100 nm) that has been explored. It is also compatible with a wide range of rigid and flexible receiver substrates, including glass, Si, plastics, and paper (Figure 2). Thus this approach is a highly scalable, low-cost, efficient method of assembling

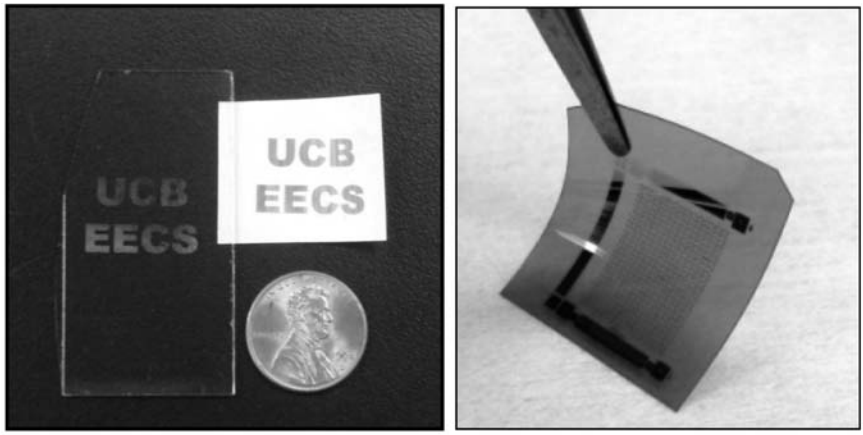


FIGURE 2 Printed NW arrays on unconventional substrates: glass and paper (left) and plastic (right). Refer to Figure 3 for high-magnification images. Source: Yerushalmi et al., 2007. Reprinted with permission.

functional NWs on substrates and may point the way toward the realization of high-performance, flexible electronics based on printed, single-crystalline, high-mobility nanoengineered materials. Notably, the printed NW arrays are highly aligned in the direction of rolling and are limited to a monolayer (Figure 3) with no uncontrolled aggregations.

To shed light on the transfer mechanism and the process dynamics, and to gain better control of the printing process, we have explored the effect on the density of printed NWs of modifying the surface chemical of the receiver substrate (Fan et al., 2008b). As shown in Figure 3b, for the $–CF_3$ terminated $SiO_2$ surfaces (which are highly hydrophobic and not sticky), we observed almost no significant transfer of NWs ($<10^{-3}$ NW/μm) from the donor to the receiver substrate. Using an identical printing process on $–NH_2$ and $–N(Me)_3^+$ terminated $SiO_2$ (which are highly hydrophilic and sticky), we observed a high-density transfer of NWs, approaching ~8 NW/μm (Fan et al., 2008b). This major modulation of printed NW density by ~4 orders of magnitude demonstrates the importance of nanoscale chemical interactions during the printing process.

A lubricant (octane and mineral oil, 2:1, *v*:*v*) is applied to all surfaces during printing. The lubricant, which serves as a spacing layer between the two substrates, minimizes NW-to-NW friction, uncontrolled breakage, and detachment of NWs. The results suggest that during the printing process NWs are dragged across a receiver substrate and are eventually detached from the roller as they are anchored to the surface-functional groups of the receiver substrate by van der Waals forces.

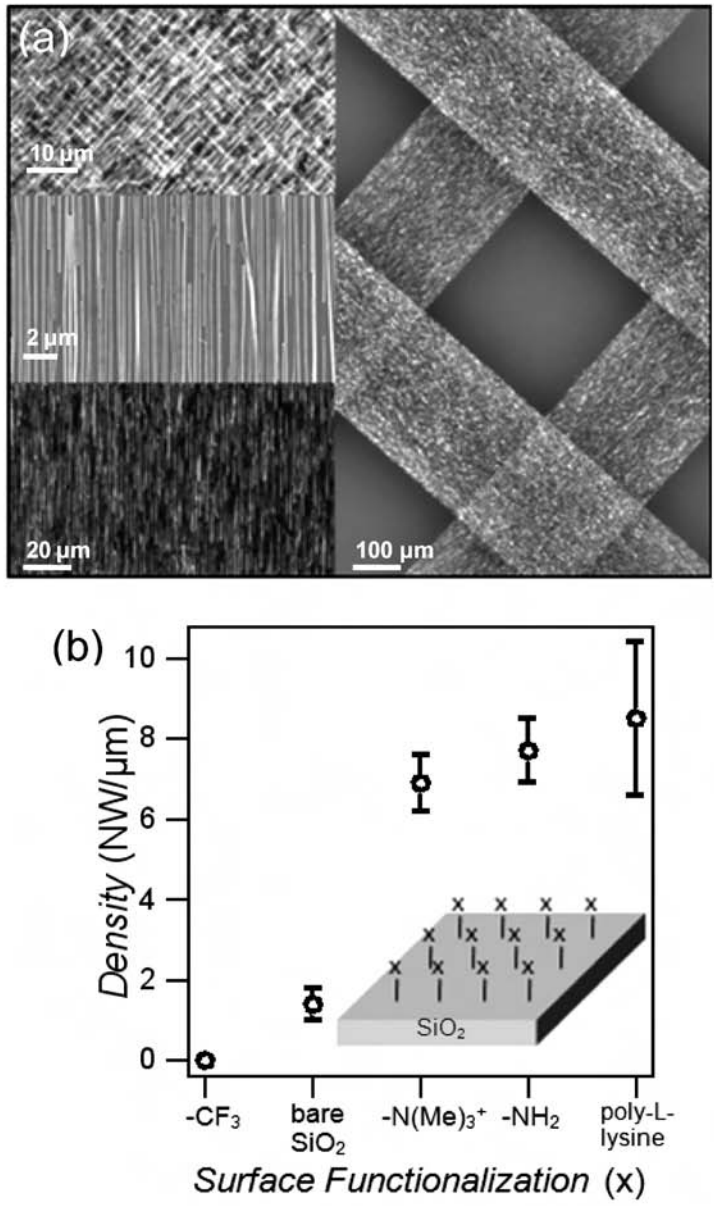


FIGURE 3 (a) Optical images of printed Ge NW arrays (middle left image is a scanning electron micrograph). The printed NWs are ~30 nm in diameter. (b) Printed nanowire density as a function of the surface functionalization of the receiver substrate. Source: Fan et al., 2008b. Copyright 2008 ACS.

## PRINTED NANOWIRE ARRAYS FOR INTEGRATION IN ELECTRONIC DEVICES

We have successfully demonstrated highly uniform assembly of parallel arrays of NWs on the wafer scale, which is crucial for the fabrication and integration of high-throughput devices (Fan et al., 2008b). After patterned printing of NW arrays on the receiver substrates, which can be crystalline Si, low-cost glass, or bendable/flexible plastic, device structures can be fabricated using conventional lithography methods, with each device consisting of a parallel array of NWs.

In the most commonly explored device configurations, metal source/drain (S/D) and gate contacts are deposited by evaporation and liftoff. Because NWs are randomly positioned, not all of the printed NWs in a given region bridge the S/D electrodes. Since there is minimal NW-to-NW crossing or bundling in our assembled NWs, only the NWs that directly bridge S/D electrodes contribute to conduction. This technology is most relevant for printable macroelectronics with channel widths on the order of tens of microns or more and does not cause large device variations or degrade performance.

By tuning the width of the patterned regions for the assembly, the on-current can be readily modulated so more NWs will be involved in conduction (Figure 4) (Fan et al., 2008b). The observed linear dependence of the on-current on the device width illustrates the uniformity and reproducibility of NW printing tech-

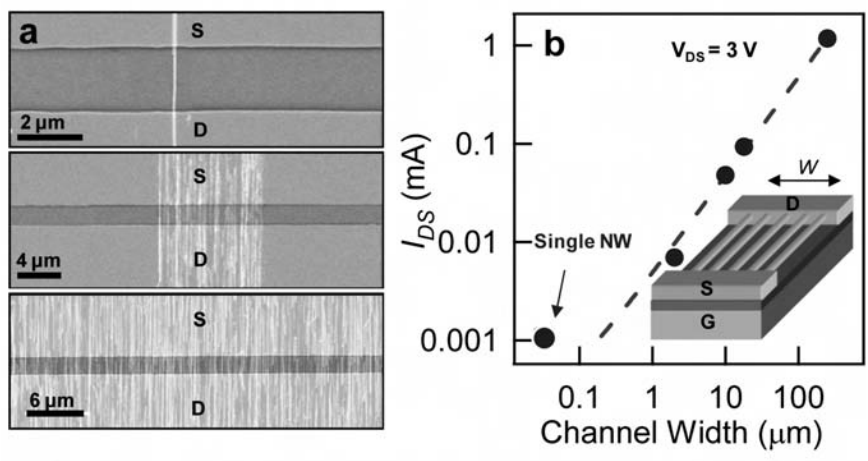


FIGURE 4 Devices based on printed NW arrays. (a) From top to bottom, scanning electron microscope images of back-gated, single GeNW FET, 10 μm and 250 μm wide, parallel arrayed NW FETs. (b) On-current as a function of channel-width scaling, showing a highly linear trend. Source: Fan et al., 2008b. Copyright 2008 ACS.

nology over large areas. Specifically, a standard deviation σ~15 percent in the on-current (for a width of ~200 μm) was observed (Javey et al., 2007).

## HETEROGENEOUS ASSEMBLY FOR INTEGRATION IN MULTIFUNCTIONAL CIRCUITS

In addition to device integration, there is a great deal of interest in the development of a versatile method of heterogeneous integration of crystalline materials on substrates to add functionality to a device (e.g., combining sensing capability with conventional electronics). Because NW printing technology is done at ambient temperatures, it is uniquely suited for the heterogeneous assembly of crystalline NWs on substrates for integration in multifunctional circuits (Fan et al., 2008a).

For instance, high-mobility Ge NWs can be printed at certain locations on the receiver substrates to enable high-performance transistors, while optically active CdSe NWs (direct band gap, Eg~1.8eV) can be printed at other predefined sites to enable efficient photo detection (Fan et al., 2008a). This is in distinct contrast to conventional Si processing for which the integration of crystalline-compound semiconductors has proven to be challenging because of lattice mismatches and interface problems.

The fabrication of heterogeneous NW circuits involves two-step printing of heterostructured Ge/Si and CdSe NWs at predefined locations on substrates, followed by device and circuit fabrication using conventional microfabrication processing. As a proof of concept of the feasibility of using NW printing technology for heterogeneous circuitry, we fabricated Ge/Si NW amplifiers and CdSe photo detectors that are integrated on-chip on Si substrates (Figure 5). The CdSe NW photo detectors were shown to be highly responsive to white light (~100x reduction in resistance upon irradiation to ~4 mW/cm$^2$), and the integrated Ge/Si NW FETs amplified the signal of the sensors by ~1000x.

For this demonstration, we fabricated large arrays of the proof-of-concept circuits on substrates; each circuit was used as an individual pixel to detect light and amplify the signal. Owing to the high uniformity and reproducibility of the printing process, a relatively large matrix (13 × 20) of the all-NW sensor circuits was fabricated on a chip (with a yield of greater than 80 percent) and used as an integrated imager (Figure 6) (Fan et al., 2008a). In the future, the yield can be significantly improved by optimizing NW synthesis and fabrication processing.

To demonstrate the imaging capability, a circular halogen light source was focused and projected onto the center of the array, and the circuit output current was measured and normalized on a 0–100 scale with "0" and "100" representing the minimum and maximum measured intensity. The output profile map clearly matches the variation in spatial intensity of the light source, with the intensity decreasing from the center to the outer edge of the circuit (Fan et al., 2008a).

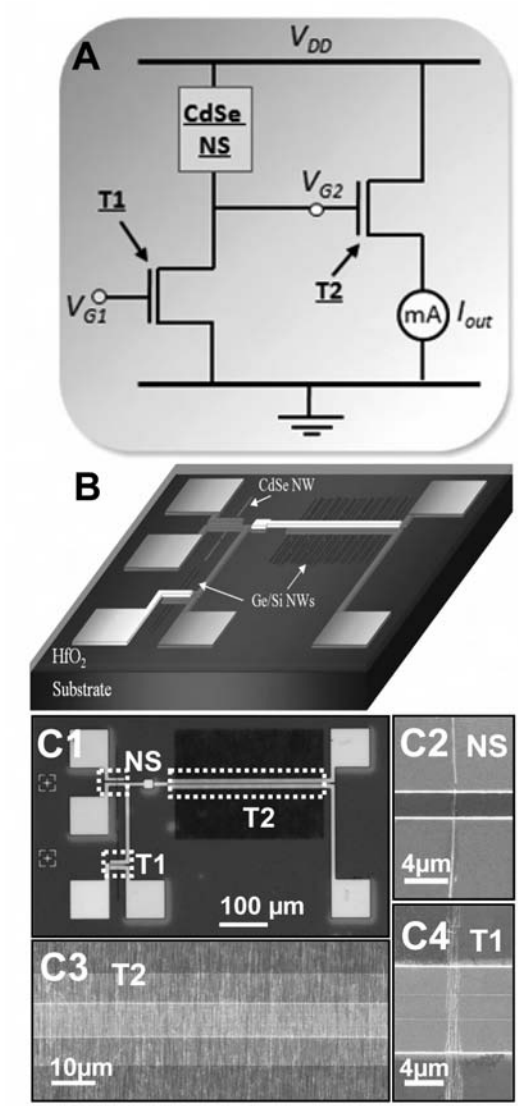


FIGURE 5 Heterogeneous NW assembly for all integrated sensor circuitry. (A) Circuit diagram for the all-NW photo detector, with high mobility Ge/Si NW FETs (T1 and T2) amplifying the photo response of a CdSe nanosensor. (B) Schematic drawing of the all-NW optical-sensor circuit based on ordered arrays of Ge/Si and CdSe NWs. (C1) An optical image of the fabricated NW circuitry, consisting of a CdSe nanosensor (NS). (C2) Two Ge/Si core/shell NW FETs (T2 and T1). (C3) and (C4) channel widths of ~300 μm and 1 μm, respectively. Each device element in the circuit can be independently studied for dynamics and circuit debugging. Source: Fan et al., 2008a. Reprinted with permission.

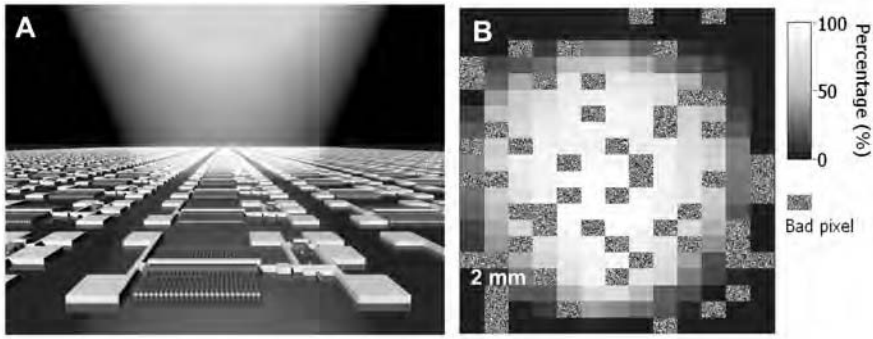


FIGURE 6 NW sensor circuitry with imaging functionality. (A) Schematic diagram. (B) An output profile of the integrated imager for a circular light spot (gray pixels represent defective sites). Source: Fan et al., 2008a. Reprinted with permission.

Each pixel size can be further down-scaled in the future by reducing the feature sizes, such as channel and interconnect lengths and widths. This work not only demonstrates NW device integration at an unprecedented scale, but also presents a novel system based on printed NW arrays that may have a number of technological applications with NWs as building blocks.

## CONCLUSION

Significant progress has been made in the roll printing of NWs for highly ordered assembly of crystalline semiconductors on foreign substrates with high uniformity, regularity, and tunable density. Parallel arrays of NWs have been shown to be high-performance building blocks for diodes, transistors, and sensors that can be readily integrated into functional circuits on unconventional substrates, such as bendable plastics. In addition, heterogeneous integration can be achieved using a multistep printing process at ambient temperatures. This approach may lead to the development of a wide range of novel printable electronics that are unattainable with conventional Si processing.

## ACKNOWLEDGMENTS

This work was supported by DARPA/MTO, Intel, and MARCO MSD. The nanowire synthesis was supported by a LDRD from Lawrence Berkeley National Laboratory. In addition, Johnny Ho has a graduate fellowship from Intel Foundation. All fabrication was performed in the Berkeley Microfabrication Laboratory.

## REFERENCES

Ahn, J.-H., H.-S. Kim, K.J. Lee, S. Jeon, S.J. Kang, Y. Sun, R.G. Nuzzo, and J.A. Rogers. 2006. Heterogeneous three-dimensional electronics by use of printed semiconductor nanomaterials. Science 314(5806): 1754–1757.

Bryllert, T., L.-E. Wernersson, L.E. Froberg, and L. Samuelson. 2006. Vertical high-mobility wrap-gated InAs nanowire transistor. IEEE Electron Device Letters 27(5): 323–325.

Fan, Z., J.C. Ho, Z.A. Jacobson, H. Razavi, and A. Javey. 2008a. Large scale, heterogeneous integration of nano-wire arrays for image sensor circuitry. Proceedings of the National Academy of Sciences of the United States of America 105(32): 11066–11070.

Fan, Z., J.C. Ho, Z.A. Jacobson, R. Yerushalmi, R.L. Alley, H. Razavi, and A. Javey. 2008b. Wafer-scale assembly of highly ordered semiconductor nanowire arrays by contact printing. Nano Letters 8(1): 20–25.

Ford, A.C., J.C. Ho, Z. Fan, O. Ergen, V. Altoe, S. Aloni, H. Razavi, and A. Javey. 2008. Synthesis, contact printing, and device characterization of ni-catalyzed, crystalline InAs nanowires. Nano Research 1(1): 32–39.

Friedman, R.S., M.C. McAlpine, D.S. Ricketts, D. Ham, and C.M. Lieber. 2005. High-speed integrated nanowire circuits. Nature 434(7037): 1085.

Huang, Y., X. Duan, Y. Cui, L.J. Lauhon, K.-H. Kim, and C.M. Lieber. 2001. Logic gates and computation from assembled nanowire building blocks. Science 294(5545): 1313–1317.

Javey, A., S. Nam, R.S. Friedman, H. Yan, and C.M. Lieber. 2007. Layer-by-layer assembly of nanowires for three-dimensional, multifunctional electronics. Nano Letters 7(3): 773.

Lee, K., M.J. Motala, M.A. Meitl, W.R. Childs, E. Menard, J.A. Rogers, R.G. Nuzzo, and A. Shim. 2005. Large area, selective transfer of microstructured silicon (us-Si): a printing-based approach to high-performance thin-film transistors supported on flexible substrates. Advanced Materials 17(19): 2332–2336.

McAlpine, M.C., R.S. Friedman, and C.M. Lieber. 2005. High-performance nanowire electronics and photonics and nanoscale patterning on flexible plastic substrates. Proceedings of the IEEE 93(7): 1357–1363.

Reuss, R.H., B.R. Chalamala, A. Moussessian, M.G. Kane, A. Kumar, D.C. Zhang, J.A. Rogers, M. Hatalis, D. Temple, G. Moddel, B.J. Eliasson, M.J. Estes, J. Kunze, E.S. Handy, E.S. Harmon, D.B. Salzman, J.M. Woodall, M.A. Alam, J.Y. Murthy, S.C. Jacobsen, M. Olivier, D. Markus, P.M. Campbell, and E. Snow. 2005. Macroelectronics: perspectives on technology and applications. Proceedings of the IEEE 93(7): 1239–1256.

Service, R.F. 2000. New age semiconductors pick up the pace. Science 287(5452): 415–417.

Someya, T., and T. Sakurai. 2003. Integration of Organic Field-Effect Transistors and Rubbery Pressure Sensors for Artificial Skin Applications. Presented at IEEE International Electron Devices Meeting, Washington, D.C., December 7–10, 2003.

Wang, X., J. Song, J. Liu, and Z.L. Wang. 2007. Direct-current nanogenerator driven by ultrasonic waves. Science 316(5821): 102–105.

Xiang, J., W. Lu., Y. Hu, Y. Wu, H. Yan, and C.M. Lieber. 2006. Ge/Si nanowire heterostructures as high performance field-effect transistors. Nature 441(7092): 489–493.

Yerushalmi, R., Z.A. Jacobson, J.C. Ho, Z. Fan, and A. Javey. 2007. Large scale, highly ordered assembly of nanowire parallel arrays by differential roll printing. Applied Physics Letters 91(20): 203104.

Yu, Y., V. Protasenko, D. Jena, H.L. Xing, and M. Kuno. 2008. Photocurrent polarization anisotropy of randomly oriented nanowire networks. Nano Letters 8(5): 1352.

Zhang, D.H., Z.Q. Liu, C. Li, T. Tang, X.L. Liu, S. Han, B. Lei, and C. Zhou. 2004. Detection of $NO_2$ down to ppb levels using single and multiple $In_2O_3$ nanowire devices. Nano Letters 4(10): 1919–1924.

Zheng, G., F. Patolsky, Y. Cui, W.U. Wang, and C.M. Lieber. 2005. Multiplexed electrical detection of cancer markers with nanowire sensor arrays. Nature Biotechnology 23(10): 1294–1301.

# The Role of DNA in Nanoarchitectonics

MIHRIMAH OZKAN
*University of California, Riverside*

CENGIZ S. OZKAN
*University of California, Riverside*

In the last several decades, the scaling of complementary metal oxide semiconductor (CMOS) technologies has fueled multiple industries, which have produced new industrial and defense products. However, the International Technology Roadmap for Semiconductors (ITRS) anticipates that scaling will necessarily end, perhaps by 2016, with a 22 nanometer (nm) pitch length (9 nm physical gate length). To address that eventuality, ITRS defines several potential avenues for research, such as bioinspired assembly, that could lead to new paradigms and alternative technologies. The ultimate goal is the development of highly controlled, high-throughput fabrication of nanoelectronics as stand-alone devices/systems or components/devices that could be integrated heterogeneously onto existing device platforms.

Deoxyribonucleic acid (DNA) and peptide nucleic acids (PNAs), which have base sequences that offer specificity, are attractive assembly linkers for bottom-up nanofabrication. Recent publications on bioassembly describe ex vivo-assembled discrete devices, such as DNA-single-walled carbon nanotubes (SWNTs) and virus-nanocrystal (NC) nanoarchitectures for electronics components (Tseng et al., 2006; Wang et al., 2006) and the programming of nucleic-acid sequences

for the large-scale assembly of nanostructures (Akin et al., 2007; Ruan et al., 2007).

We believe that novel routes, which would be available with self-assembly processing and highly integrated materials, could circumvent current challenges of CMOS to achieve environmental friendliness, thermal balance, dielectric quality, and manageable capital costs of next-generation fabrication facilities—if we can develop massively parallel integration of SWNTs and semiconducting, defect-tolerant nanowires.

Assembly based on biomolecular recognition is a promising approach for constructing complex architectures from molecular building blocks, such as SWNTs and NCs (Ravindran et al., 2003). In the Ozkans' laboratories at the University of California, Riverside (UCR), researchers are using a "tiered" approach to the nanomanufacturing of molecular electronics to address several issues: gaining an understanding of charge-carrier transport across bio-inorganic interfaces; ensuring error-free repeatability of the synthesis of hybrid building blocks; and directing the integration of nanoscale components (including assembled architectures, nanowires, and nanodevices) on silicon (Si) platforms. Figure 1 shows two

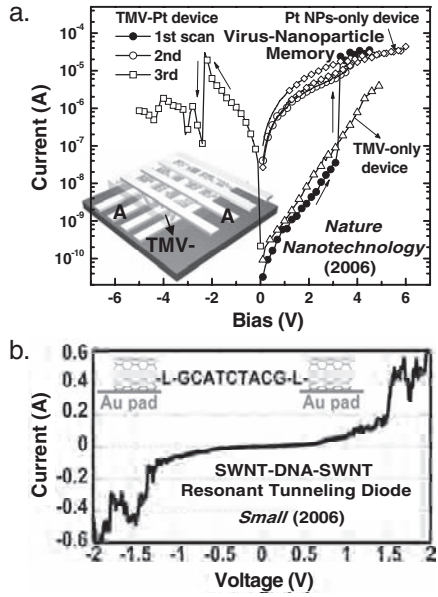


FIGURE 1 (a) Tobacco Mosaic Virus (TMV) for cross bar-memory applications. (b) DNA-CNT nanoarchitectures for resonant tunneling diodes.

novel devices fabricated at UCR: (a) a virus-NC memory device with write-erase cycles, and (b) a resonant tunneling diode based on DNA-SWNT architectures.

## CARBON NANOTUBE-BASED FUNCTIONAL NANOSTRUCTURES

The synthesis of hybrid nanoarchitectures based on SWNT-DNA or SWNT-PNA conjugates may offer unique possibilities for nanoelectronics and biotechnology (Figure 2). New structures would combine the electrical properties of SWNTs with the self-assembling properties of oligonucleotides or other biomaterials, such as proteins, enzymes, and viruses. For example, we recently demonstrated that SWNT-DNA-SWNT conjugates can be used to fabricate resonant tunneling diodes (Wang et al., 2006). Based on this result, we expect that novel devices and applications, such as bioelectronic devices, DNA sensors, mechanical actuators, templates for hierarchical assembly, and others, can be derived.

Several studies have reported using SWNTs for imaging probes in scanning force microscopy (Bernholc et al., 2002; Wong et al., 1998), and electrochemical studies have shown that SWNTs can be used as enzyme-based sensors and DNA sensors (Britto et al., 1996; Davis et al., 1997; Melle-Franco et al., 2004; Wang et al., 2004c; Zhao et al., 2002). Because SWNT electrodes have demonstrated

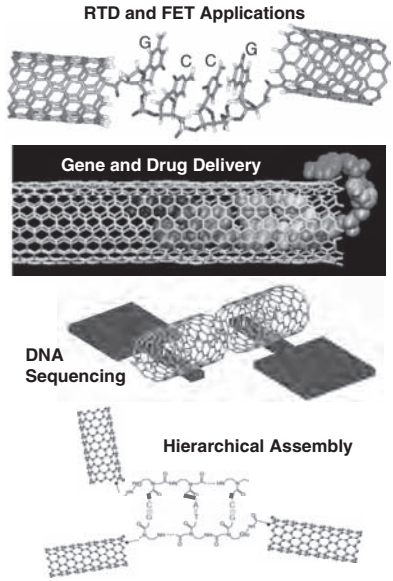


FIGURE 2 SWNT-DNA sensors for hybrid nanoelectronics, biosensors, and bottom-up nanofabrication.

catalytic properties, they could also be used as electrodes in fuel cells and electrochemical detectors in medical and military settings (Que et al., 2004; Rubianes and Rivas, 2003; Sherigara et al., 2003; Wang et al., 2004a,b; Wohlstadter et al., 2003).

Functionalized nanotubes have been used in fabricating FETs for use in nanoelectronics and biosensors (Bradley et al., 2004; Javey et al., 2003; Star et al., 2003); and several studies have shown that SWNTs and multiwalled nanotubes (MWNTs) can accommodate the encapsulation of nanoparticles, fullerenes, and metallized DNA fragments (Cui et al., 2004; Davis et al., 1998; Dennis and Briggs, 2004; Gao et al., 2003). Other studies have suggested that organic and inorganic molecules might be conjugated to the side walls of carbon nanotubes (CNTs) (Hirsch, 2002; Lin et al., 2003; Sarikaya et al., 2003; Shim et al., 2002).

## BOTTOM-UP FABRICATION: HYBRID NANOARCHITECTURES

SWNTs are being used as active components in solid-state nanoelectronics (Tsukagoshi et al., 2002), and individual SWNTs have been used to realize molecular-scale electronic devices, such as single-electron (Postma et al., 2001) and field-effect transistors (Tans et al., 1998). Several SWNT-based devices have been successfully integrated into logic circuits (Bachtold et al., 2001) and transistor arrays (Javey et al., 2002). However, the difficulty of determining the precise location and interconnection of nanotubes has so far stymied progress toward the integration of larger scale circuits.

The search for alternative routes based on molecular recognition between complementary strands of DNA has prompted an exploration of the electronic properties of DNA for use in molecular electronics and templated nanostructures (Arkin et al., 1996; Coffer et al., 1996; Heath and Ratner, 2003; Seeman, 1998, 1999, 2003). We have synthesized SWNT-DNA and SWNT-PNA conjugates, in which DNA or PNA sequences are covalently bonded to the ends of SWNTs to form a viable bio-inorganic interface (Figure 3).

Research on the fabrication of oligonucleotide-based nanoarchitectures has been focused mostly on noncovalent interactions between DNA fragments and SWNTs (Dwyer et al., 2002; Zheng et al., 2003). Because the intrinsically low conductivity of DNA limits its usefulness in electronic circuits, some investigators have attempted to distribute metal particles on the backbone of DNA to lower its resistance (Spyro, 1980; Winfree et al., 1998).

The synthesis of end-specific SWNT-DNA and SWNT-PNA complexes (Figure 3) is a novel concept that was studied for the first time at UCR (Wang et al., 2006). In the preliminary experiments, we used ssDNA with a nine-base configuration of [5'(NH2)GCATCTACG] and ssPNA with a custom sequence of (NH2)-Glu-GTGCTCATGGTG-Glu-(NH2). In order to preserve the superior electrical characteristics of SWNTs, their side walls must be free of damage or defects. Therefore, functionalization of SWNTs only at the ends, before the assembly

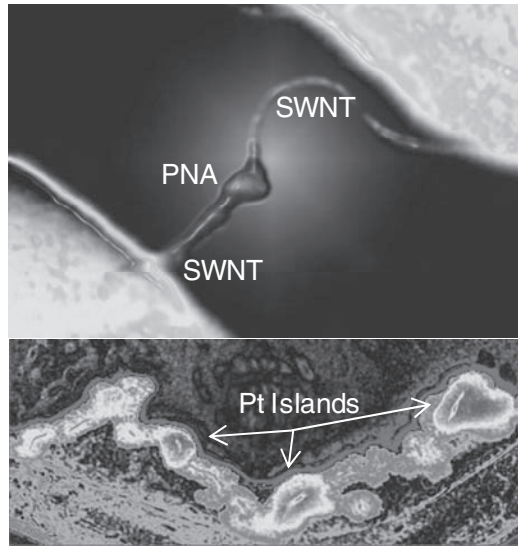


FIGURE 3 (Top) Electron microscopy image of end-to-end assembly of two SWNTs via PNA. (Bottom) Electron microscopy image of Pt metallized PNA strand. Notice formation of Pt islands during the metallization process.

process, is critical. Our work demonstrates the first successful end-to-end assembly of SWNTs using nucleic acids. After placing physical metallic contacts on SWNTs, we investigated the electrical characteristics of this heterojunction. The results show negative resonance tunneling behavior that can be adopted to fabricate resonant tunneling diode circuits.

## METALLIZED NANOARCHITECTURES

For an electrical circuit to have fast processing capability, the conductivity of circuit elements can be important. Information must be delivered to the other parts of the circuit with no delay (or loss). To achieve this, we adjusted the conductivity of the assembled circuit elements. In functional assembly such as SWNT-PNA-SWNT, the PNA link may have to be engineered to make it more conductive. We used a metallization procedure to improve the conductivity of nucleic acid-based linkers.

In one case, we developed a platinum (Pt) metallization process. The synthesis of Pt-decorated SWNT-ssDNA complexes requires a two-step chemical reduction and the deposition of metallic colloids (Mertig et al., 1998, 1999; Pompe et al., 1999; Richter et al., 2000). In the first step, SWNT-ssDNA conjugates were mixed with a salt solution (e.g., K2PtCl4 solution). After this activation step, the Pt (II)

was reduced to metallic platinum. In the reduction process, Pt dimers formed heterogeneously on DNA molecules, and the initial heterogeneous Pt nuclei quickly developed into bigger particles, consuming the metal complex feedstock in the solution (Ciacchi, 2002) to create metallized linkers (Figure 2). Because oxidized SWNTs have higher adsorption capacities for heavy metal ions (Braun et al., 1998), the Pt ions would be absorbed on SWNTs if the metallization process was done after assembly.

## MODELING OF BAND STRUCTURES AND CARRIER TRANSPORT FOR BIO-INORGANIC INTERFACES

An analysis of high-lying occupied molecular orbitals (HOMO) and low-lying unoccupied molecular orbitals (LUMO) reveals the structural and electrical properties of bio-inorganic interfaces, such as CNT/protein, quantum dot (QD)/DNA, QD/protein, metal/DNA, and metal/protein systems. In a recent study, the electrical properties of the interfaces between SWNT-ssDNA and SWNT-ssPNA were deduced via density functional theory (DFT) calculations (Singh et al., 2006; Wang et al., 2006), in which two unit cells of zigzag (10,0) oxidized CNT were linked to a DNA sequence with amine to form an amide linkage.

When the highest HOMO and lowest LUMO surface plots (shown in Figure 4) were generated, the HOMO-LUMO gap was found to be about 3.1 electron-volts (eV). For comparison, the HOMO-LUMO gap of SWNT alone is ~3.1 eV.

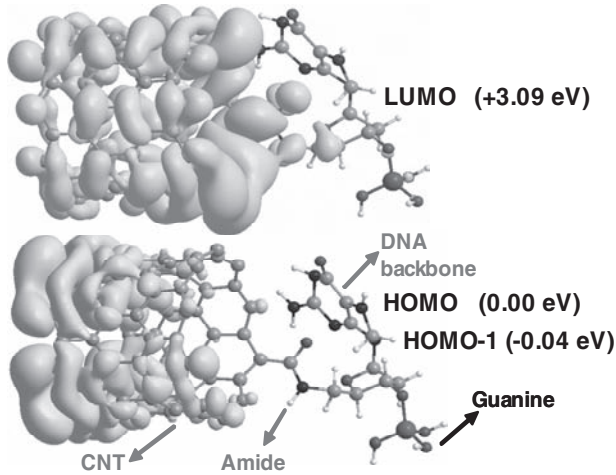


FIGURE 4 HOMO-LUMO calculation of SWNT. The gap is found to be 3.1eV. Similar modeling studies can reveal electrical characteristics of organic-inorganic interfaces.

The large gap is the result of the shortness (just two unit cells) of the modeled SWNT. For an extended (10,0) CNT, the bandgap is ~0.98 eV. The HOMO orbital is confined on the SWNT, while the LUMO orbital extends across the amide link, suggesting a good possibility of electron transfer across the amide bridge for n-type SWNTs.

Similar calculations for SWNT-ssPNA revealed that, although the HOMO orbital is confined to the glutamate link, the LUMO orbital extends over the SWNT, suggesting that SWNT-ssPNA conjugates might be used to build hole-conducting devices. Thus these preliminary studies suggest that bio-inorganic interfaces achieved by conjugating SWNTs with ssDNA and ssPNA might lead to the fabrication of n-type and p-type devices, which might someday provide an alternative or an enhancement to conventional CMOS technology.

## NANOPATTERNING VIA DIELECTROPHORESIS USING MICRO- AND NANOARRAYS

Micro- and nanoarray platforms can be used to control the electrophoretic manipulation of (bio)molecules, particles, and micro-light emitting diodes (LEDs) as electronic elements. The platform shown in Figure 5 is used for electric-field-assisted manipulation and the assembly of nanoelements, such as metallic and semiconducting SWNTs, QDs, dendrimers, and/or conjugation molecules, such as DNA fragments. The nanochip platform (shown in Figure 5) enables rapid, parallel transport within seconds to a specific location on the chip array by providing independent current or voltage control on each electrode.

Current commercialized applications of this platform include DNA hybridization and DNA analysis for molecular diagnostics via fluorescence detection using fluorophore-labeled reporters (Akin et al., 2007; Dubois and Nuzzo, 1992; Ruan et al., 2007; Salem et al., 2004). Commercial uses of DNA detection include highly multiplexed, fully validated assays and panels for identifying cystic fibrosis, respiratory viruses, hereditary hemochromatosis, and other medical conditions.

So far, different types of arrays (with 10,000, 400, and 100 sites) have been developed using silicon micromachining with fully automated and robotized fluidics. Figures 5c and 5d show the in situ assembly for the manipulation, direction, and assembly of nanoelements using electric-field assembly. The electrode array, with geometry configurable to the desired application, is energized to attract and combine different types of nanoelements (Figure 5b). When electric-field assembly is used, the process is significantly different from self-assembly in a static solution, because it enables site-specific assembly.

In the future, the controlled parallel assembly of nanowires and nanotubes could be investigated by attaching one end of a nanowire to the target DNA immobilized on the nanoarray and the other end to a reporter-DNA sequence equipped with a fluorescent tag (Figure 5d). Upon hybridization, the presence of fluorescence could be used to assess and record in situ assembly.

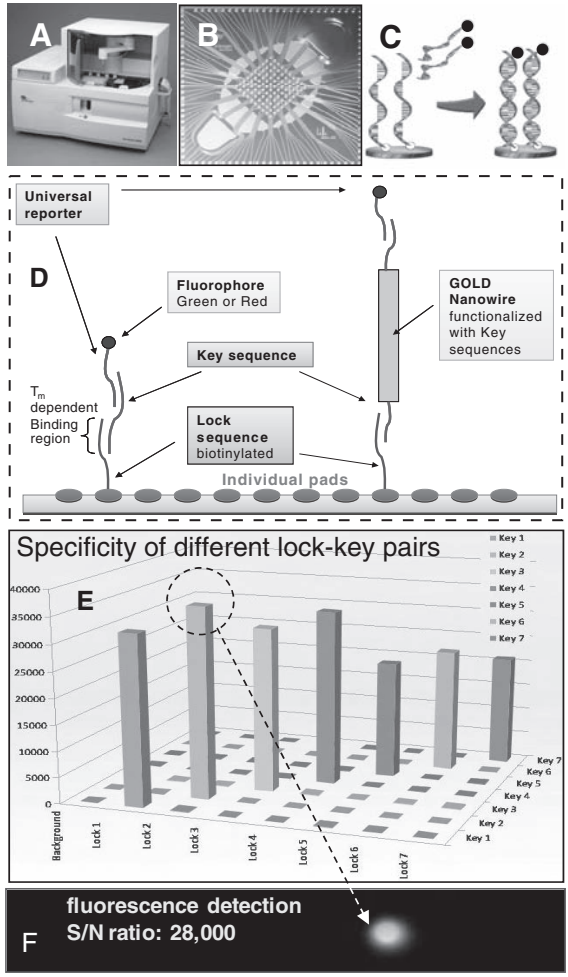


FIGURE 5 (A)–(C) Nanogen platform and microarray device for dielectrophoresis applications. (D) Assembly of ssDNA sequences and functionalized nanowires onto Si arrays. (E) Specificity of assembly of different lock and key ssDNA sequences. (F) High S/N ratio is obtained.

## CONCLUSIONS

Clearly, chemical and biological assemblies are promising technologies. However, many new technologies must be developed and much science must be learned for that promise to be fully understood and realized. We anticipate that new engineering concepts will be discovered in the near future that will enable the massively parallel assembly of nanodevices. The future of assembly engineering

(and hierarchical fabrication) may depend on being able to manipulate and control more than one type of molecular force. We anticipate that the first applications in this area will be enabled by top-down approaches for integrating assembled components onto existing device platforms.

## REFERENCES

Akin, H., J. Ruan, S. Raghunathan, X. Wang, J. Hartley, C. Ozkan, and M. Ozkan. 2007. Engineered Nucleic Acid Base Pairing for Controlled Assembly of Nanostructures. SRC TECHCON, September 2007.

Arkin, M.R., E.D.A. Stemp, R.E. Holmlin, J.K. Barton, A. Hörmann, E.J.C. Olson, and P.F. Barbara. 1996. Rates of DNA-mediated electron transfer between metallointercalators. Science 273(5274): 475–480.

Bachtold, A., P. Hadley, T. Nakanishi, and C. Dekker. 2001. Logic circuits with carbon nanotube transistors. Science 294(5545): 1317–1320.

Bernholc, J.D., M.D. Brenner, W. Nardelli, V. Meunier, and C. Roland. 2002. Mechanical and Electrical Properties of Nanotubes. Pp. 347–375 in Annual Review of Materials Research, Vol. 32. Palo Alto, Calif.: Annual Reviews.

Bradley, K., M. Briman, A. Star, and G. Grüner. 2004. Charge transfer from adsorbed proteins. Nano Letters 4(2): 253–256.

Braun, E., Y. Eichen, U. Sivan, and G. Ben-Yoseph. 1998. DNA-templated assembly and electrode attachment of a conducting silver wire. Nature 391(6669): 775–778.

Britto, P.J., K.S.V. Santhanam, and P.M. Ajayan. 1996. Carbon nanotube electrode for oxidation of dopamine. Bioelectrochemistry and Bioenergetics 41(1): 121–125.

Ciacchi, L.C. 2002. Growth of Platinum Clusters in Solution and on Biopolymers: The Microscopic Mechanisms. Ph.D dissertation. University of Dresden, Germany.

Coffer, J.L., S.R. Bigham, X. Li, Y. Rho, G. Young, R.M. Pirtle, and I.L. Pirtle. 1996. Dictation of the shape of mesoscale semiconductor nanoparticle assemblies by plasmid DNA. Applied Physics Letters 69(25): 3851–3853.

Cui, D., C.S. Ozkan, S. Ravindran, Y. Kong, and H. Gao. 2004. Encapsulation of Pt-labelled DNA molecules inside carbon nanotubes. Mechanics & Chemistry of Biosystems 1(2): 113–121.

Davis, J.J., R.J. Coles, and H. Allen, and O. Hill. 1997. Protein electrochemistry at carbon nanotube electrodes. Journal of Electroanalytical Chemistry 440(1-2): 279–282.

Davis, J.J., M.L.H. Green, H. Allen, O. Hill, Y.C. Leung, P.J. Sadler, J. Sloan, A.V. Xavier, and S.C. Tsang. 1998. The immobilisation of proteins in carbon nanotubes. Inorganica Chimica Acta 272(1-2): 261–266.

Dennis, T.J.S., and G.A.D. Briggs. 2004. Molecular motion of endohedral fullerenes in single-walled carbon nanotubes. Angewandte Chemie 116(11): 1410–1413.

Dubois, L.H., and R.G. Nuzzo. 1992. Synthesis, Structure, and Properties of Model Organic Surfaces. Pp. 437–463 in Annual Review of Physical Chemistry, Vol. 43. Palo Alto, Calif.: Annual Reviews.

Dwyer, C., M. Guthold, M. Falvo, S. Washburn, R. Superfine, and D. Erie. 2002. DNA-functionalized single-walled carbon nanotubes. Nanotechnology 13(5): 601–604.

Gao, H., Y. Kong, D. Cui, and C.S. Ozkan. 2003. Spontaneous insertion of DNA oligonucleotides into carbon nanotubes. Nano Letters 3(4): 471–473.

Heath, J.R., and M.A. Ratner. 2003. Molecular electronics. Physics Today 56(5): 43–49.

Hirsch, A. 2002. Funktionalisierung von einwandigen Kohlenstoffnanoröhren. Angewandte Chemie 114: 1933–1939. Also appeared in Angewandte Chemie International Edition 41(11): 1853–1859 with the title Functionalization of Single-Walled Carbon Nanotubes.

Javey, A., Q. Wang, A. Ural, Y. Li, and H. Dai. 2002. Carbon nanotube transistor arrays for multistage complementary logic and ring oscillators. Nano Letters 2(9): 929–932.

Javey, A., J. Guo, Q. Wang, M. Lundstrom, and H. Dai. 2003. Ballistic carbon nanotube field-effect transistors. Nature 424(6949): 654–657.

Lin, T., V. Bajpai, T. Ji, and L. Dai. 2003. Chemistry of carbon nanotubes. Australian Journal of Chemistry 56(7): 635–651.

Melle-Franco, S.M., M. Marcaccio, D. Paolucci, F. Paolucci, V. Georgakilas, D.M. Guldi, M. Prato, and F. Zerbetto. 2004. Cyclic voltammetry and bulk electronic properties of soluble carbon nanotubes. Journal of the American Chemical Society 126(6): 1646–1647.

Mertig, M., R. Kirsch, and W. Pompe. 1998. Biomolecular approach to nanotube fabrication. Applied Physics A: Materials Science & Processing 66(Suppl. 1): S723–S727.

Mertig, M., R. Kirsch, W. Pompe, and H. Engelhardt. 1999. Fabrication of highly oriented nanocluster arrays by biomolecular templating. European Physical Journal D–Atomic, Molecular, Optical and Plasma Physics 9(1-4): 45–48.

Pompe, W., M. Mertig, R. Kirsch, R. Wahl, L. Ciacchi Colombi, J. Richter, R. Seidel, and H. Vinzelberg. 1999. Formation of metallic nanostructures on biomiolecular templates. Zeitschrift für Metallkunde 90: 1085–1091.

Postma, H.W.Ch., T. Teepen, Z. Yao, M. Grifoni, and C. Dekker. 2001. Carbon nanotube single-electron transistors at room temperature. Science 293(5527): 76–79.

Que, J., Y. Shen, X. Qu, and S. Dong. 2004. Preparation of hybrid thin film modified carbon nanotubes on glassy carbon electrode and its electrocatalysis for oxygen reduction. ChemComm 1: 34–35.

Ravindran, S., S. Chaudhary, B. Colburn, M. Ozkan, and C.S. Ozkan. 2003. Covalent coupling of quantum dots to multi-walled carbon nanotubes for electronic device applications. Nano Letters 3(4): 447–453.

Richter, J., R. Seidel, R. Kirsch, M. Mertig, W. Pompe, J. Plaschke, and H.K. Schackert. 2000. Nanoscale palladium metallization of DNA. Advanced Materials 12(7): 507–510.

Ruan, J., S. Raghunathan, J. Hartley, K. Singh, H. Akin, N. Portney, and M. Ozkan, 2007. Fluorescent Tag Based Metrology for Self-Assembled Molecular Devices. Frontiers of Characterization and Metrology for Nanoelectronics. 8pp. Gaithersburg, Md.: NIST.

Rubianes, M.D., and G.A. Rivas. 2003. Carbon nanotubes paste electrode. Electrochemistry Communications 5(8): 689–694.

Salem, A.K., M. Chen, J. Hayden, K.W. Leong, and P.C. Searson. 2004. Directed assembly of multisegment Au/Pt/Au nanowires. Nano Letters 4(6): 1163–1165.

Sarikaya, M., C. Tamerler, A.K.Y. Jen, K. Schulten, and F. Baneyx. 2003. Molecular biomimetics: nanotechnology through biology. Nature Materials 2(9): 577–585.

Seeman, N.C. 1998. DNA Nanotechnology: Novel DNA Constructions. Pp. 225–248 in Annual Review of Biophysics and Biomolecular Structure, Vol. 27. Palo Alto, Calif.: Annual Reviews.

Seeman, N.C. 1999. DNA engineering and its application to nanotechnology. Trends in Biotechnology 17(11): 437–443.

Seeman, N.C. 2003. DNA in a material world. Nature 421(6921): 427–431.

Sherigara, B.S., W. Kutner, and F. D'Souza. 2003. Electrocatalytic properties and sensor applications of fullerenes and carbon nanotubes. Electroanalysis 15(9): 753–772.

Shim, M., N.W.S. Kam, R.J. Chen, Y. Li, and H. Dai. 2002. Functionalization of carbon nanotubes for biocompatibility and biomolecular recognition. Nano Letters 2(4): 285–288.

Singh, K., R. Rajeev, R. Pandey, X. Wang, R. Lake, C. Ozkan, K. Wang, and M. Ozkan. 2006. Covalent functionalization of single walled carbon nanotubes with peptide nucleic acid: nanocomponents for molecular level electronics. Carbon 44(9): 1730–1739.

Spyro, T.G., ed. 1980. Nucleic Acid: Metal Ion Interactions. New York: John Wiley and Sons.

Star, A., J.-C.P. Gabriel, K. Bradley, and G. Grüner. 2003. Electronic detection of specific protein binding using nanotube FET devices. Nano Letters 3(4): 459–463.

Tans, S.J., A.R.M. Verschueren, and C. Dekker. 1998. Room-temperature transistor based on a single carbon nanotube. Nature 393(6680): 49–52.

Tseng, R.J., C. Tsai, L. Ma, J. Ouyang, C.S. Ozkan, and Y. Yang. 2006. Digital memory device based on tobacco mosaic virus conjugated with nanoparticles. Nature Nanotechnology 1(1): 72–77.

Tsukagoshi, K., N.Yoneya, S. Uryu, Y. Aoyagi, A. Kanda, Y. Ootuka, and B.W. Alphenaar. 2002. Carbon nanotube devices for nanoelectronics. Physica B: Condensed Matter 323(1-4): 107–114.

Wang, J., G. Chen, M.P. Chatrathi, and M. Musameh. 2004a. Capillary electrophoresis microchip with a carbon nanotube-modified electrochemical detector. Analytical Chemistry 76(2): 298–302.

Wang, J., S.B. Hocevar, and B. Ogorevc. 2004b. Carbon nanotube-modified glassy carbon electrode for adsorptive stripping voltammetric detection of ultratrace levels of 2,4,6-trinitrotoluene. Electrochemistry Communications 6(2): 176–179.

Wang, J., M. Li, Z. Shi, N. Li, and Z. Gua. 2004c. Electrochemistry of DNA at single-wall carbon nanotubes. Electroanalysis 16(1): 140–144.

Wang, X., F. Liu, G.T. Senthil Andavan, X. Jing, N. Bruque, R.R. Pandey, R. Lake, M. Ozkan, K.L. Wang, and C.S. Ozkan. 2006. Carbon nanotube-DNA nanoarchitectures and electronic functionality. Small 2(11): 1356–1365.

Winfree, E., F. Liu, L.A. Wenzler, and N.C. Seeman. 1998. Design and self-assembly of two-dimensional DNA crystals. Nature 394(6693): 539–544.

Wohlstadter, J.N., J.L. Wilbur, G.B. Sigal, H.A. Biebuyck, M.A. Billadeau, L. Dong, A.B. Fischer, S.R. Gudibande, S.H. Jameison, J.H. Kenten, J. Leginus, J.K. Leland, R.J. Massey, and S.J. Wohlstadter. 2003. Carbon nanotube-based biosensor. Advanced Materials 15(14): 1184–1187.

Wong, S.S., E. Joselevich, A.T. Woolley, C.L. Cheung, and C.M. Lieber. 1998. Covalently functionalized nanotubes as nanometre-sized probes in chemistry and biology. Nature 394(6688): 52–55.

Zhao, Q., Z. Gan, and Q. Zhuang. 2002. Electrochemical sensors based on carbon nanotubes. Electroanalysis 14(23): 1609–1613.

Zheng, M., A. Jagota, E.D. Semke, B.A. Diner, R.S. Mclean, S.R. Lustig, R.E. Richardson, and N.G. Tassi. 2003. DNA-assisted dispersion and separation of carbon nanotubes. Nature Materials 2(5): 338–342.

# Cognitive Engineering

# Introduction

BARRETT S. CALDWELL
*Purdue University*
*West Lafayette, Indiana*

From power plant operations to transportation systems to health care delivery, many technology developments have been implemented without taking into account how real-world constraints would limit their effectiveness. *Set Phasers on Stun*, *The Atomic Chef*, and other books and articles have highlighted the dangers of poorly considered, designed, or implemented technologies that ultimately impair, rather than enhance, human performance. An increasingly visible and vocal new generation of cognitive engineers is addressing these issues. Cognitive engineering, as described by the Human Factors and Ergonomics Society, is focused on improving systems design and training to support human cognitive and decision-making skills, particularly in applied, naturalistic settings. Thus cognitive engineering is not about designing better brains but about designing technologies that create working situations that allow people to use their brains more effectively. The presenters at the U.S. Frontiers of Engineering session on cognitive engineering emphasized improvements in systems engineering that can enhance human performance and reduce catastrophic errors in specific application domains.

# Cognitive Engineering: It's Not What You Think

Stephanie Guerlain
*University of Virginia*
*Charlottesville, Virginia*

What is cognitive engineering? It is neither brain science nor cognitive science, nor artificial intelligence, neuro-engineering, robot design, nor perhaps a myriad of other scientific fields that may come to mind when you hear the term. Cognitive engineering does, however, include aspects of all of the fields mentioned above and many others, including psychology, anthropology, computer science, design, and systems engineering. Simply put, cognitive engineering is about the understanding and designing of systems that require human intellectual work.

Since almost every human activity involves human intellectual work, it follows that cognitive engineering can be applied to just about any human activity. I've even published a paper in a peer-reviewed scientific magazine describing the cognitive engineering aspects of riding a horse in a cross-country jumping competition (Guerlain, 2001). The article got mixed reactions from professional colleagues, some very positive and others that included the likes of, "What's next, golf?" Actually, cognitive engineering methods could be applied to understanding and improving any sporting activity and have even been applied to just the activity of *watching* a sporting competition (see, for example, White et al., 2008).

Despite these somewhat "non-engineering" applications, the field of cogni-

tive engineering has grown up around the the practical need to understand and improve quality, safety, and efficiency in high-impact, complex domains, such as aviation, medicine, and nuclear power control, where poorly designed systems can lead to major accidents. In fact, the field emerged largely in response to accidents in large, complex systems that had seemingly nothing to do with the design of those systems.

For example, let's say a highly trained cockpit crew flies a perfectly good plane into the side of a mountain. This loss of "situational awareness" has in fact happened enough times that the aviation industry has coined a term for this phenomenon called "controlled flight into terrain" (CFIT). In these cases, there was no mechanical failure and the computers and other automation, processes, and so on worked as designed. A first response as to the cause of such accidents is often "human error."

However, by studying the causes of accidents, or even by studying the day-to-day activities of people using systems designed for them, it turns out that many systems are poorly designed to begin with. The computers, automation, and other engineered processes (such as procedures, handoffs during shift changes, log-books, regulatory requirements, and other aspects of passing information among people and computers) have weak spots, and, if certain events co-occur at those points, they can, collectively, cause failure.

Ironically, people working day-to-day in such systems often see these failure modes (although they may not think of them in that way) and create "work-arounds," such as placing sticky notes to remind themselves what to do or not do, or they develop an almost "intuitive" understanding of how to react if and when things start to go wrong. These workers, who are important sources of knowledge, are often overlooked by engineers, who may not have been trained in gathering information well. This is where cognitive engineers excel.

Cognitive engineers focus not only on interviewing and observing end users, but also look at intrinsic relationships and requirements of a task. For example, for air-traffic controllers, it is a fact that multiple planes are moving at varying (but constrained) speeds and altitudes, and these facts cannot be "simplified away." But one can often take advantage of the constraints in system knowledge when designing work flow, representations, and other aspects of decision-support systems.

Cognitive engineering is a subspecialty of the broader field known as ergonomics. When most people think of ergonomics, they think of physical changes to a product or tool to make it "fit" better physically. An early example of this was the design of the Reach toothbrush (Hill and Kreifeldt, 1979), which spawned a whole new field of "toothbrush design"—"mouth-friendly" technology designed to improve the task of cleaning teeth as compared to what could be achieved with the straight-handled, rectangular-shaped toothbrush that was then the norm.

Both physical and cognitive ergonomics are important, and the same environment or system can be analyzed and improved upon from both perspectives. Table 1 provides a few examples of how a cognitive ergonomist and a physical

TABLE 1 Comparative Examples of Physical and Cognitive Ergonomics

| Human Activity | Physical Ergonomics (worker safety and risks) | Cognitive Ergonomics (process safety and risks) |
|---|---|---|
| Will sitting for 8 hours… | …cause back pain? | …cause loss of attention? |
| Will excessive noise… | …cause hearing loss? | …cause operators to miss a request? |
| Do the operator displays… | …cause eye strain? | …cause a misunderstanding of the situation? |

ergonomist might analyze the same system from two different perspectives. In general, a physical ergonomist focuses on creating an environment that is safe and does not create physical stress or difficulties for workers in that environment. Cognitive ergonomists focus on creating an environment that maintains overall process safety, for example, by minimizing the chances for human error. Both analyses are important, because improvements in either can significantly reduce downtime by reducing worker injuries, accelerate overall performance time by eliminating extraneous steps, and increase worker satisfaction.

Many people claim that ergonomics is just "common sense," but given the number of engineered systems that are designed without taking into account human capabilities and limitations and that do not truly fit task requirements, I often claim that, "Unfortunately, it's not that common." As a simple example, take the challenge of finding all apartments for rent (that allow pets) within 1 mile (e.g., 20 minutes walking distance) of a particular location. This task, which can be easily specified, is almost impossible to achieve using current search engines, not because they can not accomplish the task, but because they are not set up to run this kind of query. Thus users must endlessly search, type, click, move the mouse, zoom, scroll, page, phone, bookmark, write notes, and so on. In fact, you can imagine a well-designed system that would accept such a query and return a map and directions, price, and everything else you might want to know in one easy result that could either be printed out in a logical order for driving (or, in the case of a city, a logical bus route or walking route) or downloaded directly to a GPS system. Thus even when the technology is available, system designers may not really understand user or task requirements, thus creating a system that necessitates all sorts of workarounds and extra investment of time and effort in order to accomplish a task, thereby increasing the possibility for errors or sub-optimal solutions or, equally likely, a giving up by users because the system makes accomplishing a task too difficult.

The problem is especially prevalent in the current health care system, where so much effort is required to gather together records and relevant information for a patient (particularly one who has just moved to the area or just been admitted to an emergency room) that doctors most often rely on asking the patient for a health

history. Even if all records have been sent to a hospital, the data cannot usually be easily sorted, digested, or summarized. In a large, complex medical record, there may be pages and pages of often repeated information. What constitutes a "patient overview screen" can often be likened to a car dashboard that, instead of giving you the information you need directly while driving the car (e.g., speed, gas remaining, RPMs, "change oil" warnings etc.), gives you a set of buttons that, while driving, you could click on individually to see each of these items should you so choose. This is not an overview; this is a front-end index that requires navigating to a successive set of detail pages each and every time any information is reviewed.

Electronic medical records are just emerging, but they are, unfortunately, not patient-centered. Patients may receive health care in many different places, even if they live in only one state. However, electronic medical record systems are being implemented piecemeal and are usually only integrated within a single health care institution.

One mantra of cognitive engineering is to design for data extraction, not just data availability (Hollnagel et al., 1986). Efficient data extraction by people often means pre-organizing data and presenting it a way that lets people use their pattern-recognition skills to directly "pick up" on the answer they are seeking in an efficient, "parallel" way (e.g., more data displayed does *not* require more search time). Thus analog instruments in a car can be read quickly, because the driver only has to look at the dial to see if it is "in the red"; he or she does not have to convert a particular number into the "state" of "the car needs more gas soon." Similarly, in control rooms, operators often tend to use at least one monitor to display trends in key process parameters, because operators can interpret patterns in those trends to detect important events and then use this overview information to navigate to detailed displays as appropriate.

One area of current research in cognitive engineering is designing domain-specific overview displays that directly inform practitioners of the state of the system without requiring that they interpret information scattered across several screens (e.g., see Burns, 2000; Card et al., 1999; Cushing et al., 2006; Guerlain et al., 2002; Smoot et al., 2005). Cognitive engineers consider all inputs, outputs, and decisions that a task requires and then inform or lead design teams to ensure that they understand what should be "automated," what should be displayed, in what way, and in what order.

Cognitive engineers also analyze the design of feedback systems to human operator(s) (Norman, 1990). The autopilot system in an aircraft, for example, does not need to pull back on the yoke to make the plane go up. However, even in autopilot mode, the yoke does "pull back" for the sole purpose of providing feedback to pilots about the changing state of the airplane. The pilot can see the movement of the yoke in his or her peripheral vision while performing other tasks. The pilot does not have to focus attention on a particular dial or instrument panel

to see if a number has changed on that display. Thus when cognitive engineers design a system, they consider many kinds of feedback.

Cognitive engineers also consider the context in which a system will be used. If people work in a very noisy (e.g., industrial) environment, then relying on a beep or other sound to get their attention would not be a very good design. Other contexts to consider are the state of the people who will be using the system. How can we design a system that will accommodate all levels of potential users? Can we design the system in such a way that people will be able to use it right away and become better at using it with experience rather than relying on extensive training?

In general, design is an iterative process. Cognitive engineers focus on understanding the cognitive requirements and constraints inherent in the system, designing prototypes, testing those prototypes for usability, and iterating on the designs until production. This human-centered design process is often skipped, either because of a lack of knowledge about the cognitive engineering approach or a perceived lack of time or funding. Usability experts may be brought in after implementation of a system, but it may be difficult to make changes at that point. In fact, a usable system meets the actual requirements. Figure 1 shows how far down in the process implementation should start and how early in the process task analysis and iterative design and testing should start.

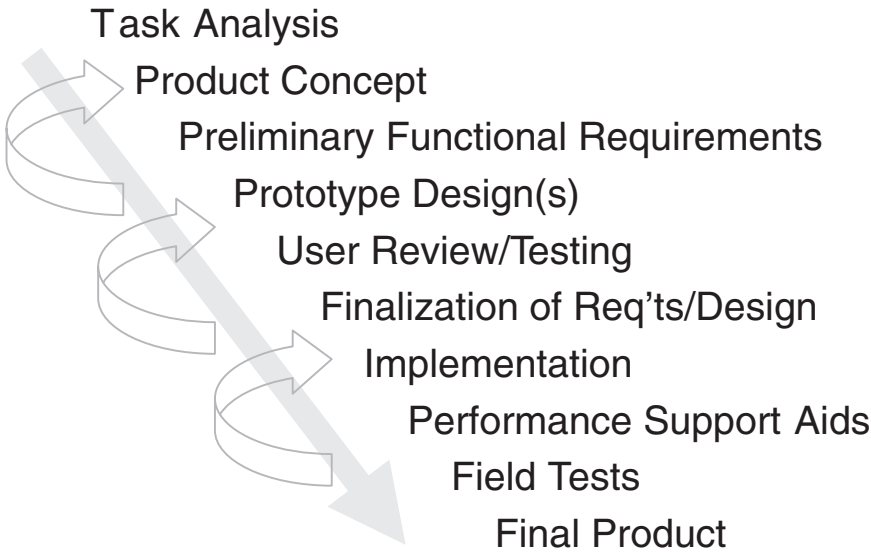


FIGURE 1 The human-centered design process.

From a practical perspective, cognitive engineering has almost limitless applications to current and evolving work practices. From a theoretical perspective, much remains to be understood about creating flexible decision-support systems that can not only support a broad range of applications, but also have the automation capabilities to put data together into a way that directly meets the task requirements.

## REFERENCES

Burns, C. 2000. Putting it all together: improving display integration in ecological displays. Human Factors 42(2): 226–241.

Card, S.K., J.D. Mackinaly, and B. Shneiderman. 1999. Readings in Information Visualization: Using Vision to Think. San Francisco, Calif.: Morgan Kaufmann.

Cushing, J., L.D. Janssen, S. Allen, and S. Guerlain. 2006. Overview + Detail in a Tomahawk mission-to-platform assignment tool: applying information visualization in support of an asset allocation planning task. Information Visualization 5: 1–14.

Guerlain, S. 2001. Judging pace cross-country: how cognitive task analysis yields insight into an apparently simple, yet complex horse-rider activity. Ergonomics in Design 9(3): 13–18.

Guerlain, S., G.A. Jamieson, P. Bullemer, and R. Blair. 2002. The MPC Elucidator: a case study in the design for human-automation interaction. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans 32(1): 25–40.

Hill, P., and J. Kreifeldt. 1979. Toothbrush—US Patent D251,038.

Hollnagel, E., G. Mancini, and D.D. Woods. 1986. Intelligent Decision Support in Process Environments. New York: Springer-Verlag.

Norman, D.A. 1990. The "Problem" with Automation: Inappropriate Feedback and Interaction, Not "Overautomation." Pp. 569–576 in Human Factors in Hazardous Situations, edited by D.E. Broadbent, A. Baddeley, and J.J. Reason. Oxford, England: Clarendon Press.

Smoot, M., E. Bass, S. Guerlain, and W. Pearson. 2005. A system for visualizing and analyzing near-optimal protein sequence alignments. Information Visualization 4(3): 224–237.

White, M., M. Odioso, M. Weaver, M. Purvis, E.J. Bass, and S. Bruce. 2008. Horses? There are horses at Foxfield? An analysis of college student hazardous drinking and related decision making behaviors. 2008 IEEE Systems and Information Engineering Design Symposium, April 25, 2008, University of Virginia, Charlottesville.

# Driving Attention: Cognitive Engineering in Designing Attractions and Distractions

*John D. Lee*
*University of Iowa*
*Iowa City, Iowa*

Driving confronts people with many of the same demands as other high-tempo, high-consequence, complex activities. People who provide health care, manage power plants, and control aircraft face similar multitasking demands, many of which are mediated by technology (Hollnagel et al., 2006; Moray, 1993; Vicente, 1999). Drivers must divide their attention among navigation, hazard detection, speed control, and lane maintenance. In addition, drivers often engage in nondriving activities, such as conversing with passengers and adjusting entertainment systems. In this multitask situation, a driver's attention is a limited, critical resource, and safety can be compromised when a driver fails to direct attention to the right place at the right time.

A recent study based on detailed data on 100 vehicles for a year showed that distractions and inattention (e.g., fatigue) contributed to approximately 80 percent of crashes and that distraction contributed to approximately 65 percent of rear-end crashes (Klauer et al., 2006). Unfortunately, this problem is likely to get worse, because driver distractions are likely to increase with rapid advances in wireless, computer, and sensor technologies (Regan et al., 2008). Not only will drivers have to manage cell phones, radios, and CD players, but they may also be tempted to use text messaging, select from MP3 music catalogs, and retrieve

information from the Internet. Rapid changes in vehicle design are being made to accommodate these new devices. Nearly 70 percent of new 2007 vehicles are compatible with MP3 players, and all 2009 Chrysler vehicles will have wireless connections to the Internet (Bensinger, 2008). These infotainment devices have the potential to make driving time more enjoyable and productive, but they also have the potential to distract drivers.

Sensor, data fusion, and control technologies promise to improve driving safety by mitigating the distraction potential of infotainment devices. Increasingly, vehicles are being equipped with sensors that monitor surrounding vehicles to identify potential collisions, warn drivers, and even respond with emergency braking. Similar technologies that can automate driving during routine situations include adaptive cruise control that accelerates and decelerates the vehicle to maintain a constant speed or constant distance from the vehicle ahead (Walker et al., 2001).

Other devices can assist drivers with emergency braking, help them keep the car centered in the lane, and attend to potential threats of collisions (Norman, 2007). Although these developments are promising, driver-support technologies may not deliver the promised safety benefits because (1) they often respond imperfectly and (2) they may encourage people to pay less attention to driving if they think the system will protect them from distraction-related lapses (Evans, 2004; Stanton et al., 1997).

As new technology has done in other domains, the introduction of infotainment and driver-support technology will fundamentally change driving. The complex array of factors that affect driving safety means that focusing simply on improving technology (e.g., designing a more capable automatic braking system) will not ensure that driving is safer, not only because technology will remain imperfect, but also because safety ultimately depends on leveraging a driver's capabilities. Technologies must be designed in a way that attracts a driver's attention to what matters most and does not annoy or distract a driver from safety-critical events.

Figure 1 illustrates the challenges of combining people and technology. The top diagram shows the complementary capacities of humans and technology—both are limited and may overlap to some degree. The middle diagram shows an effective combination of human and technological capacity—in combination, both perform better than either does alone. The bottom diagram shows a dysfunctional situation in which combined human/technology performs worse than either does alone; this can occur if the person does not capitalize on the capacity of the technology (on the left) or relies on the technology inappropriately (on the right). The disuse and/or misuse of technology often occurs when a new technology is introduced (Parasuraman and Riley, 1997). In addition, some technologies, such as warning devices, can annoy people and undermine product acceptance (on the left). Poorly coordinated technology can also interfere with a driver's ongoing response to a situation (on the right).

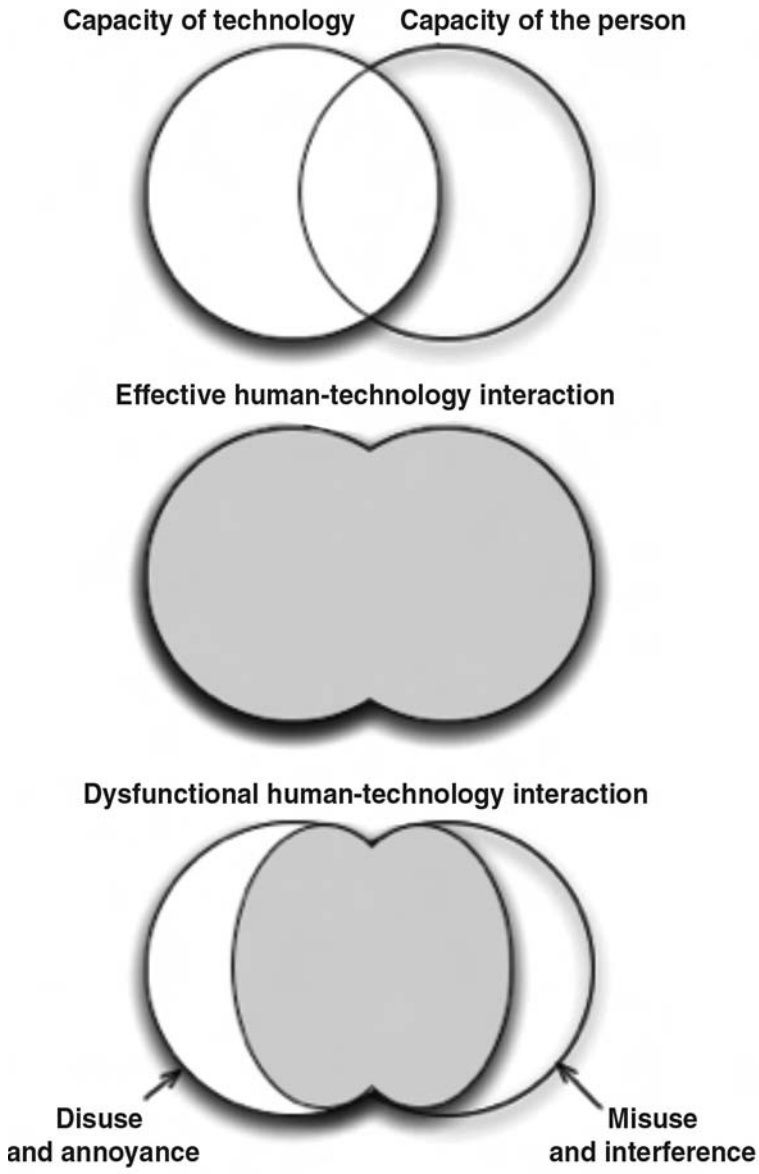


FIGURE 1 The complementary capacities of technology and humans. When properly integrated, the combination is more effective than either of them alone. When poorly integrated, the combination is less effective than either of them alone.

Achieving an effective human/technology combination requires a deep understanding of how technology mediates human attention and decision making (Lee, 2006). The dynamics of attention can be considered as a multilevel process (Michon, 1989; Sheridan, 1970). At the operational level, attention is modulated over a span of milliseconds to seconds; at the tactical level, modulation may take many seconds or minutes; and at the strategic level, it may take hours or even months. Technology can have a powerful influence at any of these levels.

Figure 2 shows the dynamics of how technology influences attention to driving and competing tasks through feed-forward, feedback, and adaptive control. With feed-forward control, drivers and technology anticipate upcoming demands and direct attention accordingly. Feedback control directs attention according to the evolving demands of the situation. Adaptive control directs attention based on changing goals and priorities. As technologies become more sophisticated and ubiquitous, they also increasingly influence drivers at all levels of attention and for each type of control.

Figure 2 shows some of the complexities associated with determining how technology mediates attention. Because both technology and humans are imperfect in directing attention to the right thing at the right time, a reliable human/technology system must perform better than either performs alone. Achieving such a design goal requires attention to the driver/technology combination rather than attention to the technology alone.

## AUGMENTATION RATHER THAN AUTOMATION

Cognitive engineering is engineering with a sensitivity to human cognitive characteristics to improve safety, performance, and satisfaction. For example, rather than using technology to automate an action in an effort to eliminate human error, a more beneficial approach, and one that may yield greater safety benefits, would be to augment, rather than automate, human capabilities.

Technology makes it possible for a vehicle to monitor both the roadway and the driver. Thus it could augment the driver's awareness of the roadway conditions and improve the driver's awareness of his or her capacity to respond to those demands. Technology might improve safety by measuring the degree to which the driver is distracted and then directing a distracted driver's attention by alerting the driver to roadway demands. In the following descriptions of how emerging vehicle technologies might mediate a driver's attention, the reader should keep in mind that similar approaches might also apply to other high-tempo, multitask activities.

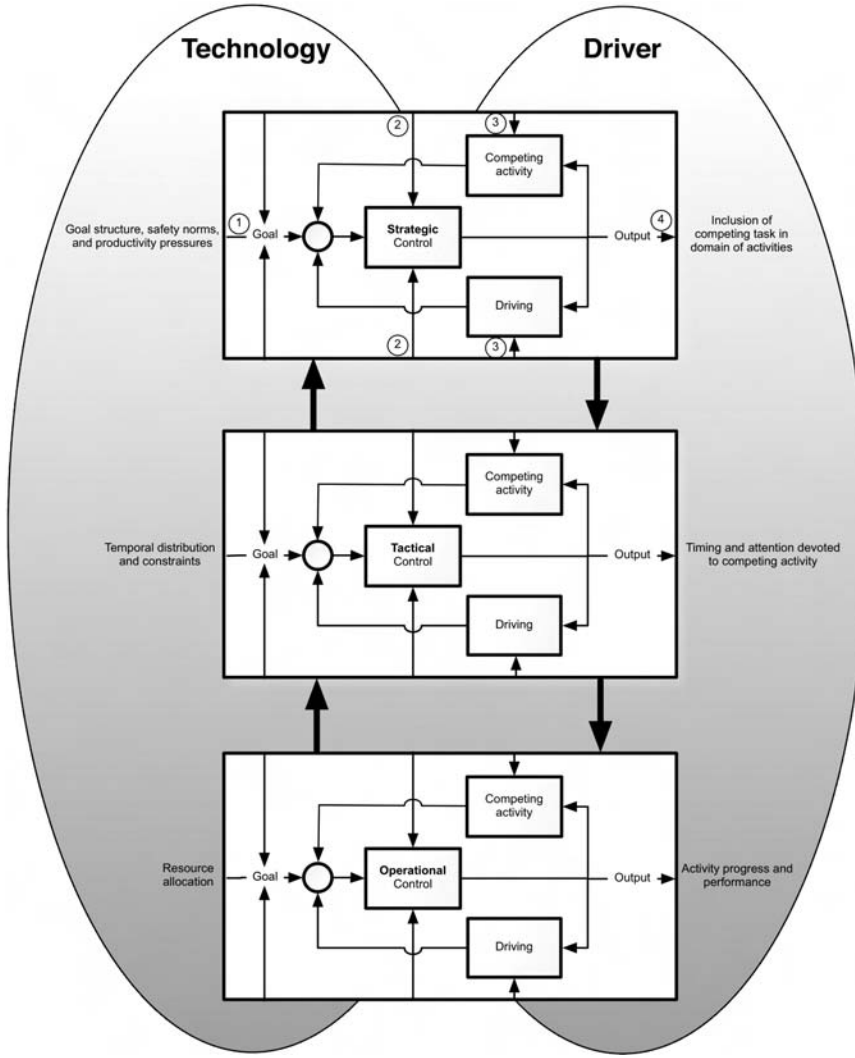


FIGURE 2 Technology-mediated attention. Numerals indicate interactions between levels: (1) adaptive control in which the output of one level affects the goal of another level; (2) feed-forward control in which the output of one level affects expectations and appropriate response schema at another level; (3) cascade effects in which the output of one level influences the control dynamics of another level; and (4) the output supports feedback control for a given level and adaptive control for other levels. The numerals at the strategic level apply also to the tactical and operational levels. The heavy lines between levels encapsulate these interactions. The ellipses in the background represent the joint control of the driver and the technology. Source: Adapted from Lee et al., 2008.

## USING MODEL-BASED DISTRACTION ESTIMATES TO IMPROVE SELF-AWARENESS

In a survey of 1,000 drivers, 80 percent said they thought they drove more safely than the average driver (Waylen et al., 2004). This sense of confidence and, perhaps, complacency is one factor that encourages drivers to divide their attention between the roadway and infotainment systems. Augmenting a driver's awareness of his or her attention to the roadway might be an effective way of mitigating distraction and helping drivers make better decisions about if and when they can safely engage in a distracting activity.

Estimating the degree of distraction experienced by a driver may be critical in helping that driver manage distraction. Figure 3 shows the output of a model of a driver switching attention between the roadway and an in-vehicle device (Hoffman, 2008). This model is based on dynamic field theory (Erlhagen and Schoner, 2002) and captures the time-varying factors that cause drivers to persist in looking away from the roadway (e.g., task inertia) and factors that draw a driver's attention back to the roadway (e.g., increasing uncertainty about the roadway situation).

The top-down, or model-driven, estimate (described above) of how drivers distribute their attention can complement a bottom-up, or data-driven, approach to estimating a driver's state based on real-time driving performance data. Bayesian networks and support vector machines are effective data-driven techniques for estimating distraction based on eye movements and steering behavior (Liang et al., 2007, in press). Increasingly instrumented vehicles provide an enormous volume of data that can be used as feedback to drivers and designers, provided those data are interpreted correctly.

Estimates of impairment related to distractions, such as text messaging, can augment a driver's awareness of impairment in three ways (Donmez et al., 2006, 2007). First, a model-based prediction of distraction could alert a driver to upcoming conflicts so that he or she can direct attention to the roadway proactively. Second, the history of distraction and the associated decrements in driving performance could be shared with drivers after a drive to help them calibrate their own estimate of how well they can manage distractions. A third approach takes into consideration the current state of the driver when redirecting his or her attention to demanding roadway situations. This approach is described in the following section.

## ALERTING AND INFORMING A DRIVER TO ENHANCE ROADWAY AWARENESS

Sensor and algorithm technologies have made it possible for a vehicle to detect hazards and alert or inform the driver, thus reducing his or her reaction time to an imminent collision (Lee et al., 2002). Unfortunately, these systems also generate many false alarms—signaling a hazard where none exists—which

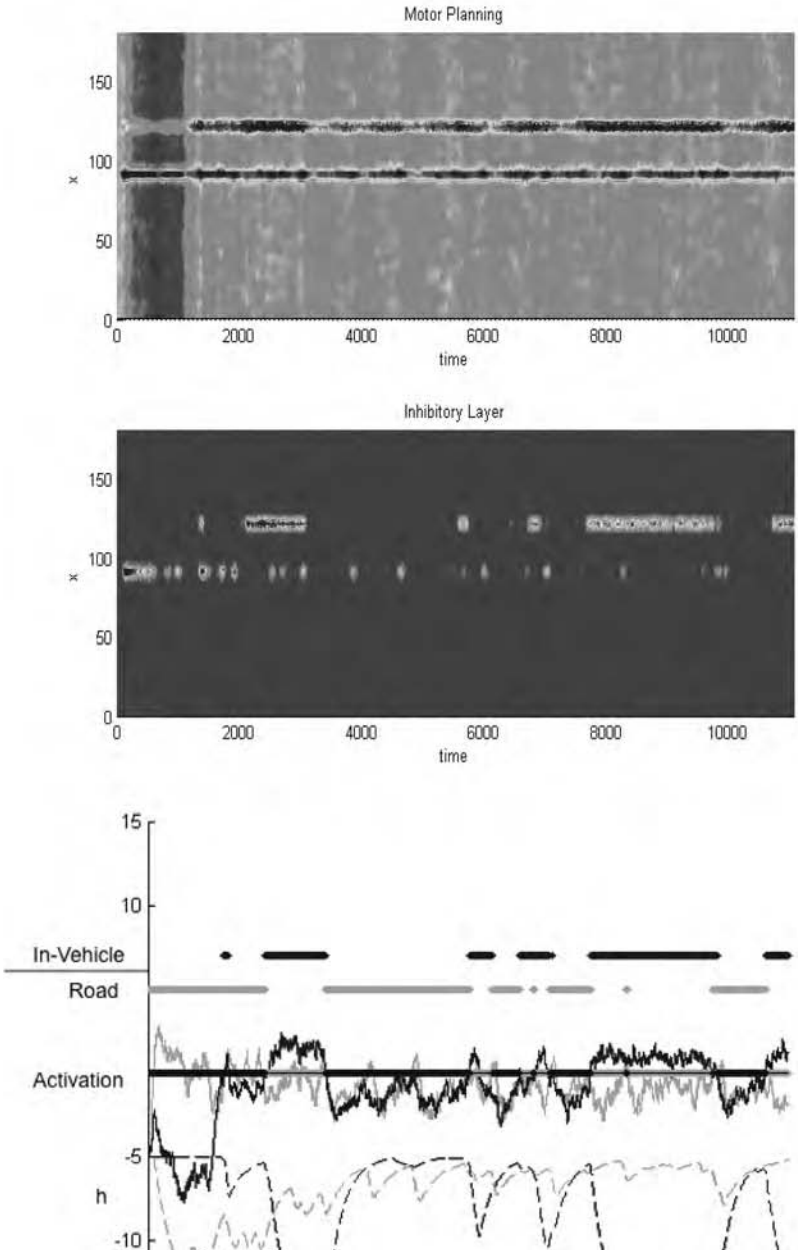


FIGURE 3 A theoretical approach to describing the dynamic distribution of attention between the roadway and an in-vehicle device.

can annoy and distract drivers. However, making such systems more useful and trusted will require more than a technological fix.

For example, based on our knowledge of human reactions, we know that drivers perceive seat vibrations as less annoying than auditory alerts (Lee et al., 2004). In addition, not all false alarms are created equal. False alarms that drivers associate with events in the environment lead them to trust the system and thus become more likely to comply with subsequent alerts. False alarms that appear as if they occur randomly tend to have the opposite effect (Lees and Lee, 2007). Drivers respond differently to alerts, even though they might all be labeled "false alarms" from a technological perspective.

Adapting a threshold for alerts based on the degree of driver distraction could reduce false alarms by raising the threshold for attentive drivers. This approach could lead to an interesting paradox in that the drivers who most need alerts are also the most likely to consider them false alerts. For example, a distracted driver might not notice a hazard (even with the alert) and so might not appreciate the value of the alert. Providing a driver with information on roadway demands and hazards after a drive, similar to the post-drive feedback for distraction, could help him or her understand the reason for the alerts. More generally, drivers are more likely to benefit from vehicle technology that augments driver attention by informing through continuous information rather than alerting through discrete warnings.

Recent studies suggest the potential benefits of post-drive feedback (McGehee et al., 2007; Tomer and Lotan, 2006). In one study, teenage drivers drove with a camera that captured abrupt braking and steering responses. The resulting video and a summary of events was shared with their parents weekly, leading to an 89 percent reduction in the number of events triggered by risky drivers compared to the baseline period. Even after the feedback was removed, the rate of events remained low until the end of the study six weeks later. Whether feedback would be accepted or effective in helping experienced drivers manage distracting technology remains to be seen.

## CONCLUSION

Technology changes the nature of driving by introducing new vulnerabilities and capacities (Woods and Dekker, 2000). Infotainment systems introduce new distractions that can undermine safety. Driver-assistance technologies promise to mitigate these distractions and improve safety. But we will not reap the potential benefits of these devices with a technology-only approach. Drivers tend to reject or misuse imperfect technologies that automate driving rather than augmenting driver capabilities. Cognitive engineering methods can show the way to using technology to leverage human capabilities to improve the safety and performance of complex systems by enhancing self-awareness and the awareness of potentially distracting technology.

Increasingly pervasive and powerful driving technologies, as in other domains, can blur the boundaries between the human and the technological, posing practical, theoretical, and philosophical issues about safety and performance, which increasingly depend on a complex interaction of driver, in-vehicle technology, and the driving situation (Lees and Lee, 2008).

Cognitive engineers face the following challenges:

- Philosophical issues relate to technologies that generally help but can also interfere with human performance. Driver-assist emergency braking, for example, generally improves crash outcomes, but, in rare instances, can impede a driver's responses.
- Practical concerns include how to draw meaning from large, complicated streams of sensor data in real time and from petabytes of accumulated data to provide feedback to operators and designers.
- Theoretical concerns relate to the dynamics of attention and how technologies can affect those dynamics and, generally, how the nature of cognition changes as technology shapes and is shaped by human activity.

## ACKNOWLEDGMENTS

Much of the research described here was part of the SAfety VEhicle(s) using the adaptive Interface Technology (SAVE-IT) Program, sponsored by the U.S. Department of Transportation–National Highway Traffic Safety Administration (NHTSA) (Project Manager: Michael Perel) and administered by the John A. Volpe National Transportation Systems Center (Project Manager: Mary D. Stearns).

## REFERENCES

Bensinger, K. 2008. Chrysler will offer wireless Internet access in 2009 models. *Los Angeles Times*. Available online at <http://www.latimes.com/business/la-fi-wificar25-2008jun25,0,1676276.story>.

Donmez, B., L.N. Boyle, and J.D. Lee. 2006. The impact of driver distraction mitigation strategies on driving performance. Human Factors 48(4): 785–804.

Donmez, B., L.N. Boyle, J.D. Lee, and G. Scott. 2007. Safety implications of providing real-time feedback to distracted drivers. Accident Analysis and Prevention 39(3): 581–590.

Erlhagen, W., and G. Schoner. 2002. Dynamic field theory of movement preparation. Psychological Review 109(3): 545–572.

Evans, L. 2004. Traffic Safety. Bloomfield Hills, Mich.: Science Serving Society.

Hoffman, J.D. 2008. A Dynamic Field Theory Model of Shared Visual Attention While Driving. Iowa City, Iowa: University of Iowa.

Hollnagel, E., D.D. Woods, and N. Leveson. 2006. Resilience Engineering: Concepts and Precepts. Burlington, Vt.: Ashgate.

Klauer, S.G., T.A. Dingus, V.L. Neale, J.D. Sudweeks, and D.J. Ramsey. 2006. The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. DOT HS 810 594. Washington, D.C.: National Highway Traffic Safety Administration.

Lee, J.D. 2006. Human Factors and Ergonomics in Automation Design. Pp. 1570–1596 in Handbook of Human Factors and Ergonomics, edited by G. Salvendy. Hoboken, N.J.: Wiley.

Lee, J.D., D.V. McGehee, T.L. Brown, and M.L. Reyes. 2002. Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Human Factors 44(2): 314–334.

Lee, J.D., J.D. Hoffman, and E. Hayes. 2004. Collision Warning Design to Mitigate Driver Distraction. Pp. 65–72 in Proceedings of CHI 2004 Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery.

Lee, J.D., M.A. Regan, and K.L. Young. 2008. Distraction as a Breakdown of Multi-level Control at Different Time Horizons. In Driver Distraction: Theory, Effects, and Mitigation, edited by M.A. Regan, J.D. Lee and K.L. Young. Boca Raton, Fla.: CRC Press.

Lees, M.N., and J.D. Lee. 2007. The influence of distraction and driving context on driver response to imperfect collision warning systems. Ergonomics 50(8): 1264–1286.

Lees, M.N., and J.D. Lee. 2008. Enhancing Information Acquisition from the Driving Environment. In Human Factors of Visual and Cognitive Performance in Driving, edited by C. Castro. Boca Raton, Fla.: Taylor and Francis.

Liang, Y., M.L. Reyes, and J.D. Lee. 2007. Real-time detection of driver cognitive distraction using support vector machines. IEEE Intelligent Transportation Systems 8(2): 340–350.

Liang, Y., M.L. Reyes, and J.D. Lee. 2007. Non-intrusive detection of driver cognitive distraction in real-time using Bayesian networks. Transportation Research Board 2018: 1–8. Washington, D.C.: The National Academies Press.

McGehee, D.V., M. Raby, C. Carney, J.D. Lee, and M.L. Reyes. 2007. Extending parental mentoring using an event-triggered video intervention in rural teen drivers. Journal of Safety Research 38(2): 215–227.

Michon, J.A. 1989. Explanatory pitfalls and rule-based driver models. Accident Analysis and Prevention 21(4): 341–353.

Moray, N. 1993. Designing for Attention. Pp. 111–134 in Attention: Selection, Awareness, and Control, edited by A. Baddeley and L. Weiskrantz. Oxford, England: Clarendon Press.

Norman, D.A. 2007. Design of Future Things. New York: Basic Books.

Parasuraman, R., and V. Riley. 1997. Humans and automation: use, misuse, disuse, abuse. Human Factors 39(2): 230–253.

Regan, M.A., J.D. Lee, and K.L. Young. 2008. Driver Distraction: Theory, Effects and Mitigation. Boca Raton, Fla.: CRC Press.

Sheridan, T.B. 1970. Big brother as driver: new demands and problems for the man at the wheel. Human Factors 12(1): 95–101.

Stanton, N.A., M. Young, and B. McCaulder. 1997. Drive-by-wire: the case of driver workload and reclaiming control with adaptive cruise control. Safety Science 27(2): 149–159.

Tomer, T., and T. Lotan. 2006. In-vehicle data recorder for evaluation of driving behavior and safety. Transportation Research Record 1953: 112–119.

Vicente, K.J. 1999. Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-based Work. Mahwah, N.J.: Lawrence Erlbaum Associates.

Walker, G.H., N.A. Stanton, and M.S. Young. 2001. Where is computing driving cars? International Journal of Human-Computer Interaction 13(2): 203–229.

Waylen, A.E., M.S. Horswill, J.L. Alexander, and F.P. McKenna. 2004. Do expert drivers have a reduced illusion of superiority? Transportation Research Part F: Psychology and Behavior 7(4-5): 323–331.

Woods, D.D., and S.W.A. Dekker. 2000. Anticipating the effects of technological change: a new era of dynamics for human factors. Theoretical Issues in Ergonomics Science 1(3): 272–282.

# Human Reliability Analysis in Cognitive Engineering and System Design[1]

RONALD LAURIDS BORING
*Sandia National Laboratories*
*Albuquerque, New Mexico*

Human factors engineering (HFE) combines elements of several engineering disciplines, psychology, and computer science into a single discipline (Boring, 2002). Two major subdisciplines of HFE include:

- *cognitive engineering* (CE), which focuses on the cognitive aspects of human-system interactions to maximize system usability (Nielsen, 1993), safety (Palanque et al., 2007), and user enjoyment (Norman, 2002)
- *human reliability analysis* (HRA), typically part of an overall probabilistic risk assessment (PRA), which focuses primarily on verifying the safe performance of human actions

Despite similarities in focus, the main difference between CE and HRA is in the timing of when they are used. CE is typically implemented in the design phase of

[1]The submitted manuscript has been authored by a contractor of the U.S. government under contract No. DE-AC04-94AL85000. The U.S. government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. government purposes.

the engineering cycle, whereas HRA is often used in the verification and validation phase, after systems have already been built. However, the application of HRA primarily to as-built systems is a historical artifact.

Analysts have included assessments of human reliability in military system evaluations since the 1960s (Swain, 1963), but the first widely publicly available guidance for HRA was described in the WASH-1400 report (U.S. Nuclear Regulatory Commission, 1975), which addresses the safety of nuclear power plants. The Technique for Human Error-Rate Prediction (THERP) HRA method (Swain and Guttman, 1983) provided the first systematic method of identifying, modeling, and quantifying human errors.

THERP and subsequent HRA methods developed in the aftermath of the Three Mile Island nuclear incident in the United States were accompanied by a call for risk-informed decision making using PRA and HRA (Kadak and Matsuo, 2007). Together, HRA and PRA produced assessments of existing systems with less emphasis on design than was typical with HFE and CE.

## HUMAN RELIABILITY PROCESS MODEL

The three phases of contemporary HRA methods are depicted in Figure 1. As shown, HRAs can be characterized as qualitative or quantitative. A qualitative HRA includes the identification and modeling phases described below. It converges on other assessment approaches such as root-cause analysis, which is used to determine the causes of human errors. A subsequent quantitative HRA uses these qualitative insights to estimate the likelihood of these errors.

### *HRA Phase 1: Identify the Sources of Errors*

This phase typically consists of a task analysis to determine human actions and a review of those actions to identify opportunities for errors. *Performance-*

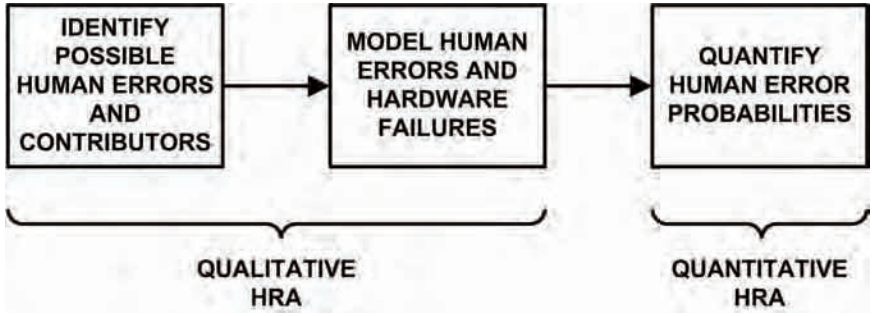


FIGURE 1 The three phases of HRA.

TABLE 1 Performance-Shaping Factors in *Good Practices for Implementing HRA*

| | | |
|---|---|---|
| Applicability and suitability of training and experience | Workload, time pressure, and stress | Accessibility or operability of equipment |
| Suitability of relevant procedures and administrative controls | Team and crew dynamics | Need for special tools |
| Availability and understandability of instrumentation | Available staffing and resources | Communications strategy and coordination |
| Time available vs. time required | Ergonomic quality of human-system interface | Special fitness needs |
| Complexity of required diagnosis and response | Environment | Off-normal operations and situation |

Source: U.S. Nuclear Regulatory Commission, 2005.

*shaping factors* (PSFs), aspects of behavior and context that may impact the outcome of a task, are then identified. For example, a PSF might be the presence or absence of clearly defined, well-understood procedures, which can greatly enhance or hinder human performance of a given task.

*Good Practices for Implementing HRA,* a report sponsored by the U.S. Nuclear Regulatory Commission (2005), provides a standardized list of 15 PSFs believed to have an impact on human performance in the nuclear domain (see Table 1). An individual HRA method may have as few as three PSFs (Galyean, 2006) or as many as 50 PSFs (Chang and Mosleh, 2007), depending on the level of detail required for capturing human actions.

### *HRA Phase 2: Model the Errors in an Overall Risk Model*

Human activities of interest in an HRA are not generally performed in isolation; they are interactions with hardware systems. The hardware systems modeled in a PRA feature reliability curves for both systems and components to provide mean times before failure. A failed hardware system can cause humans to fail at their prescribed tasks, or a human error can cause a hardware system to fail prematurely or unexpectedly.

A hardware system may be designed as a failsafe backup for human actions errors, such as an automatic pressure-venting valve that can mitigate system damage if the human operator fails to regulate pressure properly. Conversely, the human operator may save a failed hardware system. For example, positive human intervention can recover a failure or prevent the escalation of a hardware failure.

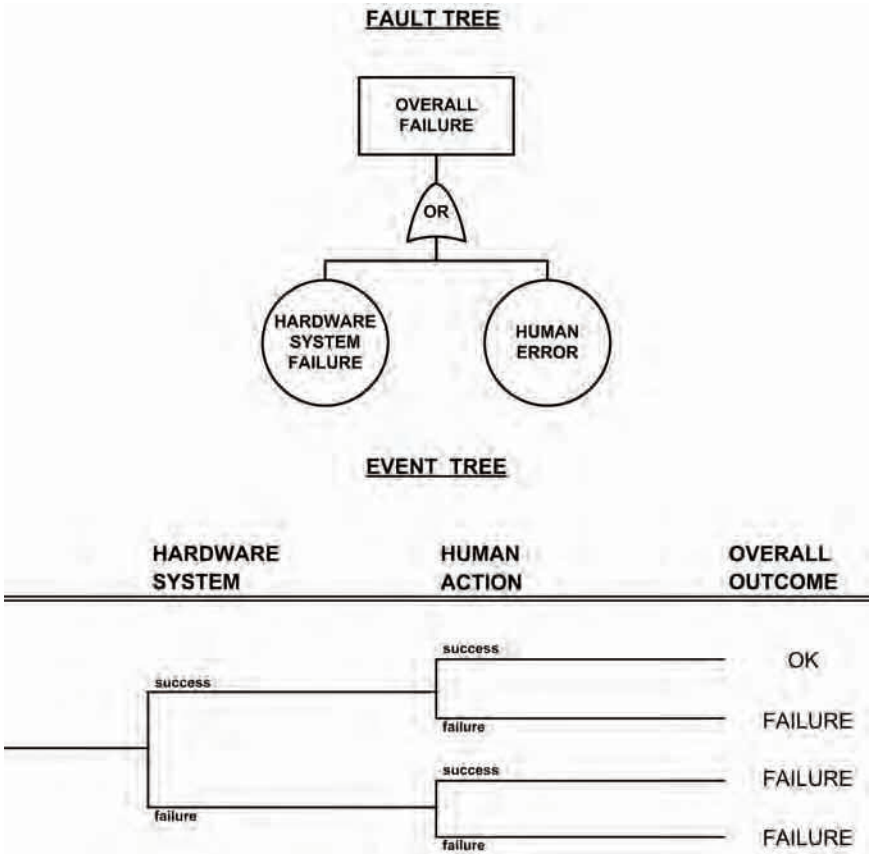


FIGURE 2 A logical "OR" gate connecting hardware-system failure and human error in the form of a fault tree (top) and event tree (bottom). The fault tree is read from bottom to top. The event tree is read as a sequence from left to right.

In an HRA, human activities are modeled as part of a fault tree, or event tree (see Figure 2), to show their interactions with the hardware system.

### *Phase 3: Quantify the Errors*

The object of many HRAs is to provide a probabilistic expression of the likelihood of a failed human action, called the *human error probability* (HEP). HRAs are primarily differentiated by their approaches to error quantification. Although dozens of approaches have been developed, they tend to follow a common pattern, beginning with a *nominal HEP* (i.e., a generic or default error rate for human

$$HEP_{overall} = HEP_{nominal} \times PSF \begin{cases} 0 < PSF < 1 \Rightarrow HEP_{overall} < HEP_{nominal} \therefore \text{reliability increases} \\ PSF = 1 \Rightarrow HEP_{overall} = HEP_{nominal} \therefore \text{reliability stays same} \\ PSF > 1 \Rightarrow HEP_{overall} > HEP_{nominal} \therefore \text{reliability decreases} \end{cases}$$

EQUATION 1

activities) and followed by a modification of the nominal HEP according to the specific PSFs.

PSFs are often treated as multipliers. For example, if the effect of good procedures has a PSF value less than one, the product of the nominal HEP and the PSF multiplier would be less than the nominal HEP, resulting in an overall decrease in HEP and corresponding increase in human reliability. Conversely, if the effect of poor procedures has a PSF value greater than one, the product of the nominal HEP and the PSF multiplier would be greater than the nominal HEP, resulting in an overall increase in HEP and corresponding decrease in human reliability (see Equation 1).

## APPLICATION OF HUMAN RELIABILITY ANALYSIS TO SYSTEM DESIGN

HRAs can be either retrospective or prospective. The purpose of a *retrospective HRA* is to assess the risk of something that has already happened, such as an incident or accident, to determine the likelihood of it happening the way it actually did. Was it an anomalous accident, or is it to be expected that it could occur again, given the same situation? A *prospective HRA* is an attempt to assess the risk of something that hasn't actually happened, such as an extremely rare event (e.g., human performance in a nuclear power plant control room during a seismic event or fire).

Note that, even though a prospective HRA can be extremely helpful for anticipating breakdowns in the human-system interface, prospective HRAs have not commonly been used to provide information that can be incorporated into the early-stage design of a system. Rather, as noted in Hirschberg (2004), prospective HRAs are usually used to improve existing processes and systems by pinpointing weaknesses and providing a basis for prioritizing "fixes." Thus, they are typically used in assessing and making iterative improvements in existing technologies.

This after-the-fact use of prospective HRAs is artificially limiting. If they were used not just on as-built systems but also on systems that are still being designed, they could be design tools used in combination with CE and HFE. Three recent developments show how HRAs could be used in the design phase of system development.

## The Need for Human-Certified, Safety-Critical Systems

Recent regulatory guidance documents, such as the *Human Factors Engineering Program Review Model* (O'Hara et al., 2004) for nuclear power plants and *Human-Rating Requirements* (NASA, 2005) for aerospace systems, suggest using HRAs as part of the design process to complement existing human-factors design best practices (Boring, 2007a). As new nuclear power and aerospace systems are built, qualitative HRAs can complement other HFE and CE techniques to anticipate sources of human errors and, ultimately, to help design the system to prevent those errors from occurring. In addition, quantitative HRAs may be used to help determine the likelihood and consequences of specific errors and to prioritize the error-likely design issues according to their impact on safety.

## The Emergence of Resilience Engineering

A recent development is a growing awareness that the negative consequences of an incident can be greatly mitigated by the quality of underlying human interactions with the system. The goal of *resilience engineering* (Hollnagel, 2006; Sheridan, 2008) is to identify the qualities that make humans, processes, and systems robust or resilient in the face of adverse events. Resilience engineering differs from HRA in that it argues for the unpredictability of adverse events, but it shares many conceptual underpinnings with HRA.

Resilience engineering can be reconciled with HRA in the context of system design. HRA provides a standardized way of assessing vulnerabilities in human actions, which make actions less robust. An HRA can even be used to define the characteristics of resilience (e.g., PSFs that characterize resilient, as opposed to brittle, actions). In the context of system design, the goals of resilience engineering and HRA are complementary, and HRA can help identify and build resilient processes and systems.

## Development of Human Reliability for Modeling Human Performance

Cacciabue (1998) and others (e.g., Boring, 2007b; Lüdke, 2005) have explained the importance of the simulation and modeling of human performance for HRA. In human-performance modeling, a virtual human (in the form of a cognitive simulation) interacts with virtual systems to reveal areas where human performance is degraded or enhanced in human-system interactions. Simulations address the dynamic nature of human performance in a way that has not been possible with classic static HRA methods.

A chief advantage of incorporating HRA into human-performance modeling is that it provides a way of estimating the safety of novel equipment and configurations. It is reasonable to assume there will also be significant cost advantages to using modeling to screen new equipment virtually instead of configuring a

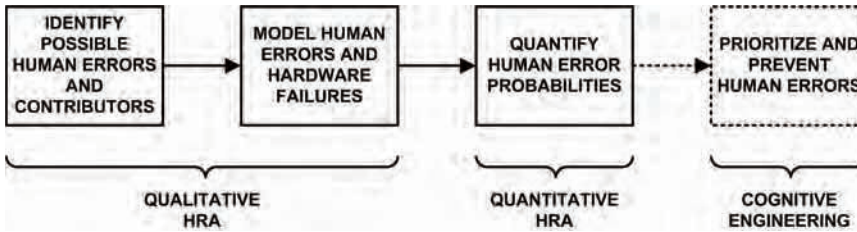


FIGURE 3 The four phases of HRA integrated with CE.

simulator with new equipment and enlisting appropriate personnel (e.g., control room staff) to perform representative tasks (Boring et al., 2008).

Human-performance modeling, utilizing insights from CE to provide a reasonable and reliable simulation, has already been shown to be a powerful system-design tool in HFE (Foyle and Hooey, 2007). When elements of HRA (such as dynamically assigned PSFs) are included in human-performance modeling, simulations can not only show if humans will interact successfully with a system, but can also provide a basis for determining the performance decrements and enhancements for particular system configurations.

## CONCLUSION

In this brief paper I have outlined the three process phases typically associated with HRA, namely identification, modeling, and quantification. These three phases represent a historic evolution that should now evolve to include a fourth phase, error prevention, particularly in the design phase of systems (see Figure 3).

Insights based on 25 years of experience with formal HRAs can now be applied to a process more closely aligned with HFE and CE. Insights derived from HRAs on the types and causes of human errors, as well as the likelihood and consequences of those errors, will ultimately facilitate the design of safer systems.

## REFERENCES

Boring, R.L. 2002. Human-computer Interaction as Cognitive Science. Pp. 1767–1771 in Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Boring, R.L. 2007a. Meeting Human Reliability Requirements through Human Factors Design, Testing, and Modeling. Pp. 3–8 in Risk, Reliability and Societal Safety. Vol. 1: Specialisation Topics. Proceedings of the European Safety and Reliability Conference (ESREL 2007), edited by T. Aven and J.E. Vinnem. London, U.K.: Taylor and Francis.

Boring, R.L. 2007b. Dynamic Human Reliability Analysis: Benefits and Challenges of Simulating Human Performance. Pp. 1043–1049 in Risk, Reliability and Societal Safety. Vol. 2: Thematic Topics. Proceedings of the European Safety and Reliability Conference (ESREL 2007), edited by T. Aven and J.E. Vinnem. London, U.K.: Taylor and Francis.

Boring, R.L., D.I. Gertman, T.Q. Tran, and B.F. Gore. 2008. Framework and Application for Modeling Control Room Crew Performance at Nuclear Power Plants. Pp. 930-934 in Proceedings of the 52nd Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Cacciabue, P.C. 1998. Modelling and simulation of human behaviour for safety analysis and control of complex systems. Safety Science 28: 97–110.

Chang, Y.H.J., and A. Mosleh. 2007. Cognitive modeling and dynamic probabilisitic simulation of operating crew response to complex system accidents. Part 2: IDAC performance influencing factors model. Reliability Engineering and System Safety 29: 1014–1040.

Foyle, D.C., and B.L. Hooey. 2007. Human Performance Modeling in Aviation. Boca Raton, Fla.: CRC Press.

Galyean, W.J. 2006. Orthogonal PSF Taxonomy for Human Reliability Analysis. Pp. 1–5 in Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, May 14–18, 2006, New Orleans, Louisiana. Paper PSAM-0281. Washington, D.C.: American Society of Mechanical Engineers.

Hirschberg, S. 2004. Human Reliability Analysis in Probabilistic Safety Assessment for Nuclear Power Plants. CSNI Technical Opinion Papers No. 4**,** OECD NEA No. 5068. Issy-les-Moulineaux, France: OECD Nuclear Energy Agency.

Hollnagel, E. 2006. Resilience—The Challenge of the Unstable. Pp. 9–17 in Resilience Engineering: Concepts and Precepts, edited by E. Hollnagel, D.D. Woods, and N. Leveson. Burlington, Vt.: Ashgate Publishing Company.

Kadak, A.C., and T. Matsuo. 2007. The nuclear industry's transition to risk-informed regulation and operation in the United States. Reliability Engineering and System Safety 92: 609–618.

Lüdke, A. 2005. Kognitive Analyse formaler sicherheitskritischer Steuerungssysteme auf Basis eines integrierten Mensch-Maschine-Models. Berlin: Akademische Verlagsgesellschaft Aka GmbH.

NASA (National Aeronautics and Space Administration). 2005. Human-Rating Requirements for Space Systems. NPR 8705.2A. Washington, D.C.: NASA Office of Safety and Mission Assurance.

Nielsen, J. 1993. Usability Engineering. San Francisco, Calif.: Morgan Kaufman.

Norman, D.A. 2002. Emotion and design: attractive things work better. Interactions Magazine 9(4): 36–42.

O'Hara, J.M., J.C. Higgins, J.J. Persensky, P.M. Lewis, and J.P. Bongarra. 2004. Human Factors Engineering Program Review Model. NUREG-0711, Rev. 2. Washington, D.C.: U.S. Nuclear Regulatory Commission.

Palanque, P., S. Basnyat, A. Blandford, R. Bernhaupt, R. Boring, C. Johnson, and P. Johnson. 2007. Beyond Usability for Safety Critical Systems: How to Be SURE (Safe, Usable, Reliable, and Evolvable)? Pp. 2133–2136 in CHI 2007 Conference Proceedings, Extended Abstracts. New York: Association for Computing Machinery.

Sheridan, T.B. 2008. Risk, human error, and system resilience: fundamental ideas. Human Factors 50(3): 418–426.

Swain, A.D. 1963. A Method for Performing a Human Factors Reliability Analysis. Monograph SCR-686. Albuquerque, N.M.: Sandia National Laboratories.

Swain, A.D., and H.E. Guttman. 1983. Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications (THERP). Final Report. NUREG/CR-1278. Washington, D.C.: U.S. Nuclear Regulatory Commission.

U.S. Nuclear Regulatory Commission. 1975. Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants. WASH-1400, NUREG-75/014. Washington, D.C.: U.S. Nuclear Regulatory Commission.

U.S. Nuclear Regulatory Commission. 2005. Good Practices for Implementing Human Reliability Analysis (HRA). NUREG-1792. Washington, D.C.: U.S. Nuclear Regulatory Commission.

# Cognitive Engineering Applications in Health Care

Ann M. Bisantz
*University at Buffalo,*
*State University of New York*

The goal of cognitive engineering is to support the cognitive activities associated with behavior, particularly in complex working environments, through the design of system components, such as user interfaces, automation, decision aids, and training. Health care is an environment with classic complexities—time pressure, risk, uncertainties, and many interacting components. The health care environment is further complicated by multiple levels or domains of concern. For instance, even an individual patient consists of numerous, interacting systems that may not all be well understood and for which only limited or indirect information may be available.

The complexity of the patient domain is compounded by the complex sociotechnical working environment that addresses the patient's needs—the health care system—which is comprised of many people working both individually and in teams, who must coordinate their actions and who have different, sometimes competing goals (e.g., health care providers vs. government regulators vs. insurance companies vs. hospital administrators). In the health care environment, individuals interact with a variety of information sources and technologies, ranging from handwritten charts to pagers and phones to electronic medical records and digital imaging systems. Resources in the health care environment, such as caregiver time

and hospital beds, are limited, and demands on the system (i.e., incoming patients and their conditions) are unpredictable.

Methods in cognitive engineering have been developed to uncover and represent both complexities in high-consequence fields such as health care and the knowledge and strategies experienced practitioners use to perform successfully (Bisantz and Burns, 2008; Bisantz and Roth, 2008; Crandall et al., 2006; Vicente, 1999). The results of cognitive engineering analyses can have a critical impact on the design of information, tasks, and training that will enhance, rather than disrupt, successful work practices and allow practitioners to respond appropriately to diverse, unpredictable events.

Cognitive engineering research in health care environments, which has a general goal of supporting safe and effective performance, has followed different research paths, including (1) characterizing complexities in the environment and demands on practitioners, sometimes with a focus on preventing medical errors; and (2) focusing on the design and/or impacts of new technologies. Understanding demands on practitioners, the strategies they use to meet those demands, and the role of information from different sources and technologies in work practice is essential to designing new information systems that can improve patient care.

## CHARACTERIZING COMPLEXITY: SYSTEM STRUCTURE, STRATEGIES, AND COMMUNICATION

A common method of representing the complexities of the work domain (i.e., the abstraction hierarchy, see Rasmussen et al., 1994; Vicente, 1999) is to represent high-level goals, balances and priorities, processes, and physical structures. In the individual patient system, for instance, researchers have modeled physiological functions and anatomical structures, as well as methods of controlling them, to support diagnostic decision making, understand information needs among clinicians, and design monitoring displays (Hajdukiewicz et al., 1998; Miller, 2004; Sharp and Helmicki, 1998; Watson and Sanderson, 2007).

Enomoto et al. (2006) and Burns et al. (2008) conducted a study of the tasks of cardiac-care telehealth nurses, as well as the underlying patient structure and processes, to identify the challenges they faced and the strategies they used in diagnosing cardiac patients based on phone interviews. Various innovative visualizations were designed and tested, alternately emphasizing mapping symptoms to diagnoses, clusters of co-occurring symptoms, and symptom severity. Hall et al. (2006) used similar techniques to simultaneously represent aspects of a surgical team, the patient, and the equipment used to compare problem-solving strategies used by anesthesiologists.

A particular complexity of interest in medicine is the need for multiple individuals (e.g., physicians, nurses, technicians, support staff) to communicate with each other to coordinate patient care, particularly in hospital settings. Poor

communication has been cited, for example, as a frequent cause of errors in the administration of medications (c.f. Rogers et al., 2004). Numerous cognitive engineering-oriented studies in medical environments have been conducted on communication functions, patterns, and sometimes breakdowns.

For example, Fairbanks et al. (2007) described aspects of communication, such as the type of partner, communication mode (e.g., face-to-face, phone), duration, and location of communication in a hospital emergency department (ED). They shadowed 20 caregivers (including attending physicians, residents, ED nurses, and charge nurses) to construct networks showing the communication pathways radiating from, and connecting, caregivers. Results provided insights into typical patterns of communication and the individuals or positions that were key communication nodes in the ED. For instance, nurses played a central role in communication; most communication was face-to face; and overall, there were frequent communications of short duration.

Potential gaps in information flow were also identified. For instance, triage nurses and ambulance personnel (emergency medical services [EMS]), who have initial contact with patients, were observed to communicate primarily with charge nurses (responsible for workflow and patient assignment) but not with the physicians who would care for the patients. This gap may indicate an opportunity for intervention, such as a change in training or procedures or the development of new technologies, such as real-time patient information systems that can be accessed by both EMS and ED staff.

Similarly, Moss et al. (2002) characterized the mode, recipient, and topic of communications by an operating room charge nurse responsible for coordinating patient, surgical team, equipment, and room preparation; the goal of the study was to suggest how electronic scheduling systems could be shared and used effectively. Guerlain et al. (2007) found that training surgeons in specific types of communication and teamwork skills, such as methods of conducting pre-operative briefings, improved communication.

Several studies have investigated communication strategies during shift changes and other transitions, when one group of caregivers must transfer information about patient status to another (Nemeth et al., 2006; Patterson et al., 2005; Sharit et al., 2005; Wears et al., 2003). Patterson et al. (2005) observed nurses during shift changes in acute-care units to identify the strategies and technologies they used to obtain necessary information. Audiotaped and face-to-face communications led to different strategies. For instance, if the information was audiotaped, incoming staff could not directly question outgoing staff; however, incoming nurses tended to listen to audiotaped information as a group and talk about the status of patients, which could result in a shared awareness of patient states and team coordination to meet patients' needs.

Wears et al. (2003) contrasted two transitions between ED physicians. In one, the transition was the source of error recovery because incoming physicians suggested an alternative, ultimately correct, diagnosis. In the second, poor com-

munication was the source of a breakdown because critical information about the state of a medication order was misunderstood, and an essential treatment was delayed.

## NEW TECHNOLOGIES AND UNINTENDED CONSEQUENCES

Advanced technology has often been advocated as a way to reduce errors and adverse events in health care (Aspden et al., 2004; IOM, 1999, 2001). In many cases, however, new technologies are designed without an in-depth understanding of the work they need to support, or they are designed to address functions other than patient care (e.g., record keeping, billing). Unless the designers understand how new technologies will be used in practice and are aware of potential barriers to their use, these technologies can lead to unanticipated, undesirable consequences (Ash et al., 2004, 2007; Bisantz and Wears, 2008; Webster and Cao, 2006), such as increased workload (because of the need for new processes or workarounds to integrate them into the workflow), or serious safety compromises (if new systems are bypassed or abandoned or if critical tasks are interrupted).

For instance, in a study of new operating room technology that integrated multiple monitoring systems into a single electronic display, Cook and Woods (1996) found that the change forced practitioners to adapt their activities, as well as some aspects of the new system, to ensure that the critical information was displayed at appropriate times.

In another case, Patterson et al. (2002, 2006) studied unanticipated effects and workarounds developed after the implementation of a system intended to reduce errors by using bar codes on medications and patient wristbands to confirm the type, dosage, and timing of medication administration. Unanticipated effects included fewer physician reviews of current medications, because it was more difficult for them to access information in the computerized system than in the old paper record; and nurses feeling pressured to administer medication "on time," even when other higher priority tasks were necessary (both of which increased the chances of adverse events).

A key workaround was that nurses would type a patient's bar code number into the system or scan a secondary wristband kept separate from the patient to save time and avoid several problems. First, the cart with the scanner was difficult to maneuver, and in some cases a computer had to be plugged in to maintain battery life. Second, they no longer had to disturb sleeping patients. Finally, scanning the second wristband was often more reliable than scanning the wristband on the patient, especially for long-term patients whose wristbands had become worn or smudged. In addition, nurses could "pre-pour" medications (place medications in cups for many patients at once, rather than scanning a wristband, scanning and administering medication(s), and moving to the next patient), to increase efficiency. Scanning medications in batches also made it more likely that medications

were recorded as administered "on time" (which eliminated the work associated with documenting late medications).

In the end, although the bar code system could reduce the chances that the wrong type or dosage of medication would be chosen, the workarounds could increase the chances of a medication being given to the wrong patient. The researchers suggested both changes in the system design (e.g., simplifying the system interface; using wireless or easily maneuverable scanners; and using longer lasting computer batteries) and changes in procedures (e.g., using more realistic times for medication administration) that could reduce the likelihood of unanticipated effects or workarounds that would increase the chances of errors.

## LEARNING FROM EXISTING TOOLS AND TECHNOLOGIES

Understanding how extant tools and artifacts work in a system is a critical step in designing new systems to support the functional purposes of an artifact, rather than merely duplicating its surface features (Nemeth, 2004; Pennathur et al., 2007; Xiao, 2005). Bauer et al. (2006) conducted a detailed analysis of an artifact used in intensive care to inform the design of an electronic system. The artifact, a patient flow sheet, is a paper form that accommodates both structured and unstructured data capture (e.g., grids for sequential vital signs and free-form notes). By observing the flow sheet in use, they were able to identify the characteristics that had to be included in an electronic system.

Some features may not have been included if the new system had simply duplicated the surface features of the form. For instance, the paper form allowed information to be entered flexibly, rather than sequentially, allowed unstructured annotations (e.g., information did not have to be entered in a particular place or with keyboard characters), and allowed users to leave information out (for a discussion of the functionality of paper artifacts, see Sellen and Harper, 2003). The paper form also supported work because it was portable, grouped information in ways that allowed comparisons to be made easily, allowed flexible annotations to accommodate unique circumstances, and allowed data to be represented in familiar notation.

An electronic system could provide additional functionality, such as automated data analysis and calculations, and could give multiple caregivers access to the information at the same time. However, the new technology still had to support flexibility in annotation and commonly used notations and comparisons.

Some of our own work has focused on the implementation of new technologies in hospital emergency rooms (Pennathur et al., 2007, 2008a,b; Wears et al., 2005), where electronic patient-tracking systems are replacing manual status boards ("whiteboards"). Manual status boards (see Figure 1) provide medical and logistical information about patients and information about patient status (e.g., designated providers, treatment status, test and laboratory results, location), as well as higher level information about hospital states (e.g., number of patients

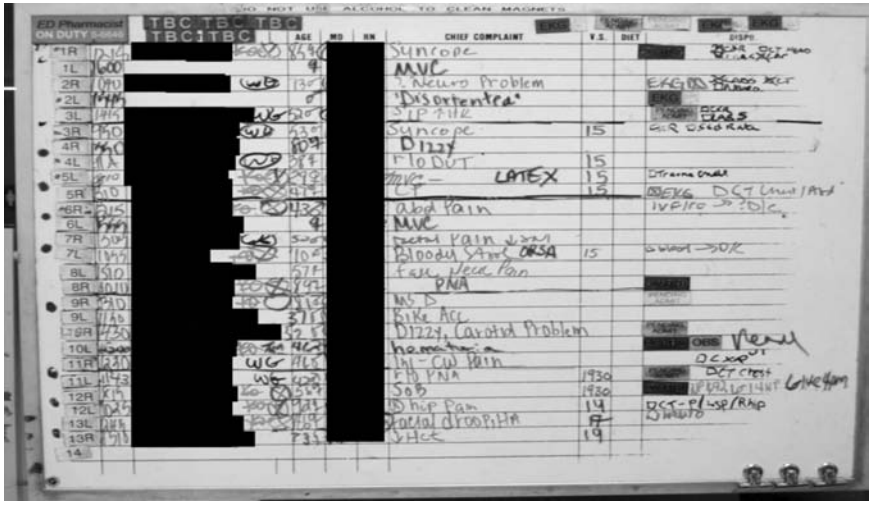


FIGURE 1 Manual whiteboard with the names of patients and providers obscured. Reprinted from *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2008*. Reprinted with permission.

in the ED, admitted patients still in the ED, available ED beds, rooms that need cleaning) and team coordination information (e.g., assignments of providers to patients or bed zones; status of on-call providers). Information on whiteboards is encoded in locally developed (e.g., by providers in the hospital or department) and locally meaningful ways. Whiteboards are used to track the process of patient care through annotations that indicate potential diagnoses, progress through treatment plans, the need for consultations or tests, and admission or discharge processes. Typically, they are located in central areas of the ED so that information is available to all care providers and can be used to coordinate activities across individuals and time (Figure 2).

Electronic status boards may mimic the look and layout of manual boards (see Figure 3), support automated recording keeping and reporting, and allow information to be accessed at different locations in the hospital, but they also impose new constraints. The ability to add or change information is limited by available computer terminals, which typically require sign-on sequences; the form of information is limited to the characters or icons available on a keyboard or through the interface, and local methods of encoding are often lost; and the length and placement of entries is prescribed (e.g., free-form annotations cannot be added).

We studied the transition from manual to electronic status boards in two university-affiliated, urban hospital EDs (Pennathur et al., 2007, 2008b; Wears

FIGURE 2 A whiteboard being viewed by multiple providers in an ED.

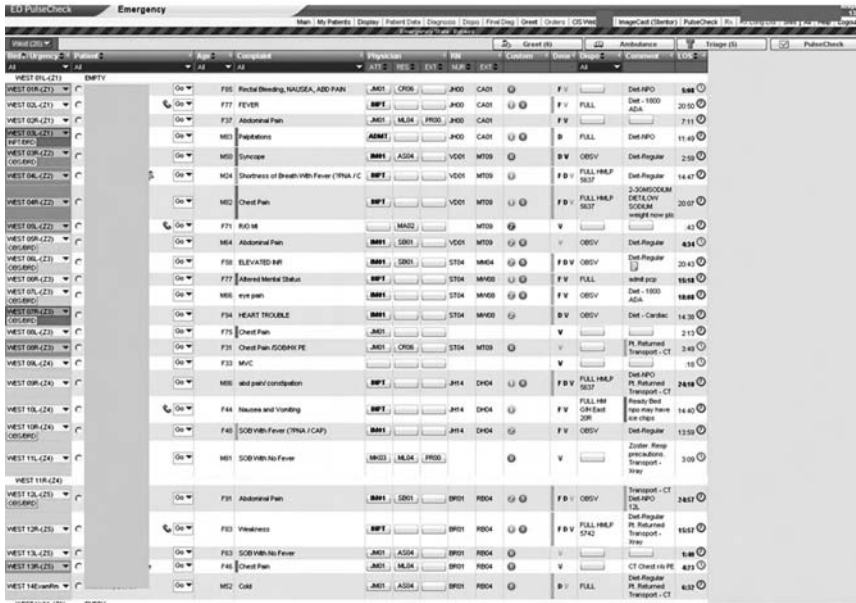


FIGURE 3 Electronic patient-tracking system screen. Source: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2008*. Reprinted with permission.

and Perry, 2007; Wears et al., 2005). One hospital had made the transition 10 months prior to our study but had continued to use manual boards along with the new system. We studied the second hospital before and after the transition. In this hospital, the manual boards were removed and replaced with the electronic system. We conducted a combination of semistructured interviews, focus groups, and observations with care providers, secretaries, information technology specialists, and administrators. We also took photographs or screen shots of the status boards at one hospital, so we could make detailed comparisons of the content and form of information in both systems.

The results of our studies indicated a number of problems related to the transition to a new technology. Shortly after the electronic system was implemented at the second hospital, providers felt that the change had a negative impact on communication and their ability to "make sense" of the overall state of the ED, in part because the system could only be viewed on desktop screens, which had limited room for displaying information and limited flexibility for documenting information about treatment plans and diagnoses. For instance, a limited number of entries were visible in the column showing treatment plans, and providers could no longer use hand-drawn checkboxes to indicate progress. Because it was more difficult for providers to document and track patient progress, some providers resorted to carrying notes; this supported the work of individual providers, but the information was no longer publicly available, thus decreasing support for coordination among caregivers.

The staff also found an unanticipated use for the system—tracking patients' dietary needs and providing a printed list of diets to the meal-delivery staff. Although this function provided a benefit to some caregivers/staff, the constraints on space in the area where dietary information was entered meant that others could not use that space to display critical clinical information (e.g., lab values). In fact, at the first hospital, where both electronic and manual boards were used, clinicians tended to rely on the manual boards, while nonclinical staff used the electronic system for administrative functions, such as finding patients or assessing room status.

Some of these difficulties could be traced to the particular implementation and interface for the system, but others were more fundamental (e.g., the removal of a public, easily modified information source that supported relatively simple coordination for each individual and among individuals).

We subsequently decided to investigate the impact of electronic patient-tracking systems on caregivers' understanding of the overall ED state, as well as specific patient information. We developed a simulation-based tracking system that allows system parameters to be varied and tested by ED staff in a laboratory setting (Pennathur et al., 2008a). Immersive, simulated environments like this are used by cognitive engineers in many domains, such as aviation and driving, to test the impact of technology designs, situations, and tasks on human operators' activities and performance (Lee et al., 2002; Sarter and Woods, 2000).

The tracking simulation we developed is based on a discrete-event simulation model of a real hospital ED and incorporates both clinical information and operational information that can be used by study participants. Historic data on patient volume and the severity of their medical conditions were used to develop the model.

This model was used to generate sets of patients with medical conditions of different levels of severity, process events (e.g., waiting, registration, triage, caregiver visits, and laboratory tests), and the duration of those events. The simulated information was augmented with demographic information, medical complaints, and time-indexed medical information (e.g., tests, results, admission decisions, and the resulting information that would be shown on a whiteboard) to create "scripts" for each simulated patient.

Different scenarios were created based on different levels of demand for ED services. The scenarios were used as input to a patient-tracking display that was created for use by participants during experiments. The scenarios were augmented with secondary tasks (e.g., phone calls or pages that had to be answered) and simulation-freeze techniques for measuring participants' awareness of information represented in the system (Endsley, 1995).

This integrated experimental system can be used to test the impact of different display-related variables (e.g., display size, mode, and format of information); operational parameters (e.g., type of caregiver, number of patients); operational tasks (use of overall monitoring and monitoring during care transitions, such as a shift change); or how ED personnel interact with and interpret information on the electronic system.

## CONCLUSION

The health care system has critical needs for improvements in efficiency, effectiveness, and safety. To meet those needs, we must first understand the complexities faced by health care workers and the knowledge, strategies, and tools they use. Cognitive engineering provides methods and tools for developing and implementing new technologies for this environment.

## ACKNOWLEDGMENTS

Funding for the studies of ED patient-tracking systems has been provided by the Emergency Medicine Foundation and the Agency for Research on Health Care Quality (grant number 1 U18 HS016672).

## REFERENCES

Ash, J.S., M. Berg, and E.W. Coiera. 2004. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. Journal of the American Medical Informatics Association 11(2): 103–112.

Ash, J.S., D.F. Sittig, R.H. Dykstra, K. Guappone, J.D. Carpenter, and V. Sehsadri. 2007. Categorizing the unintended sociotechnical consequences of computerized provider order entry. International Journal of Medical Informatics 76(Suppl. 1): S21–S27.

Aspden, P., J.M. Corrigan, J. Wolcott, and S.M. Erickson, eds. 2004. Patient Safety: Achieving a New Standard for Care. Washington, D.C.: The National Academies Press.

Bauer, D., S.A. Guerlain, and P.J. Brown. 2006. Evaluating the Use of Flowsheets in Pediatric Intensive Care to Inform Design. Pp. 1054–1058 in Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Bisantz, A.M., and C.M. Burns, eds. 2008. Applications of Cognitive Work Analysis. Boca Raton, Fla.: CRC Press.

Bisantz, A.M., and E.M. Roth. 2008. Analysis of Cognitive Work. Pp. 1–32 in Reviews of Human Factors and Ergonomics, Vol. 3, edited by D.A. Boehm-Davis. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Bisantz, A.M., and W.L. Wears. 2008. Forcing functions: the need for restraint. Annals of Emergency Medicine, forthcoming.

Burns, C.M., Y. Enomoto, and K. Momtahan. 2008. A Cognitive Work Analysis of Cardiac Care Nurses Performing Teletriage. In Advances in Cognitive Work Analysis, edited by A.M. Bisantz and C.M. Burns. Boca Raton, Fla.: Taylor and Francis.

Cook, R.I., and D.D. Woods. 1996. Adapting to new technology in the operating room. Human Factors 38(4): 593–613.

Crandall, B., G.A. Klein, and R.R. Hoffman. 2006. Working Minds: A Practitioner's Guide to Cognitive Task Analysis. Cambridge, Mass.: The MIT Press.

Endsley, M. 1995. Measurement of situation awareness in dynamic systems. Human Factors 37(1): 65–84.

Enomoto, Y., C.M. Burns, K. Momtahan, and W. Caves. 2006. Effects of Visualization Tools on Cardiac Telephone Consultation Process. Pp. 1044–1048 in Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Fairbanks, R.J., A.M. Bisantz, and M. Sunm. 2007. Emergency department communication links and patterns. Annals of Emergency Medicine 50(4): 396–406.

Guerlain, S.A., B.E. Turrentine, D.T. Bauer, J.F. Calland, and R. Adams. 2007. Crew resource management training for surgeons: feasibility and impact. Cognition, Technology & Work. Available online at <http://www.springerlink.com/content/a831373358120000/fulltext.pdf>.

Hajdukiewicz, J., D.J. Doyle, P. Milgram, K.J. Vicente, and C.M. Burns. 1998. A Work Domain Analysis of Patient Monitoring in the Operating Room. Pp. 1034–1042 in Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Hall, T.J., J.W. Rudolph, and C.G.L. Cao. 2006. Fixation and Attention Allocation and Anesthesiology Crisis Management: An Abstraction Hierarchy Perspective. Pp. 1064–1067 in Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

IOM (Institute of Medicine). 1999. To Err Is Human: Building a Safer Health System, edited by L.T. Kohn, J.M. Corrigan, and M.S. Donaldson. Washington, D.C.: National Academy Press.

IOM. 2001. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C.: National Academy Press.

Lee, J.D., D.V. McGehee, T. Brown, and M. Reyes. 2002. Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high fidelity driving simulator. Human Factors 44(2): 314–334.

Miller, A. 2004. A work domain analysis framework for modelling intensive care unit patients. Cognition Technology and Work 6(4): 207–222.

Moss, J., Y. Xiao, and S. Zubaidah. 2002. The operating room charge nurse: coordinator and communicator. Journal of the American Medical Informatics Association 9(6 Suppl. 1): S70–S74.

Nemeth, C.P. 2004. Using cognitive artifacts to understand distributed cognition. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans 34(6): 726–735.

Nemeth, C.P., J. Kowalsky, M. Brandwijk, M. Kahana, P.A. Klock, and R.I. Cook. 2006. Before I Forget; How Clinicians Cope with Uncertainty through ICU Sign-Outs. Pp. 939–943 in Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Patterson, E.S., R.I. Cook, and M.L. Render. 2002. Improving patient safety by identifying side effects from introducing bar coding in medication administration. Journal of the American Medical Informatics Association 9(5): 540–553.

Patterson, E.S., E.M. Roth, and M.L. Render. 2005. Handoffs During Nursing Shift Changes in Acute Care. Pp. 1057–1061 in Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Patterson, E.S., M. Rogers, R.J. Chapman, and M.L. Render. 2006. Compliance with intended use of bar code medication administration in acute and long-term care: an observational study. Human Factors 48(1): 15–22.

Pennathur, P., A.M. Bisantz, R.J. Fairbanks, S. Perry, F. Zwemer, and R.L. Wears. 2007. Assessing the Impact of Computerization on Work Practice: Information Technology in Emergency Departments. Pp. 377–381 in Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Pennathur, P., D. Cao, A.M. Bisantz, L. Lin, J.L. Brown, R.J. Fairbanks, T.K. Guarrera, S. Perry, and R.L. Wears. 2008a. A Simulation Study of Patient Tracking Systems. Presentation at the International Conference on Healthcare Systems, Ergonomics, and Patient Safety (HEPS). Strasbourg, France, June 25–27, 2008.

Pennathur, P., T.K. Guarrera, A.M. Bisantz, R.J. Fairbanks, S. Perry, and R.L. Wears. 2008b. Cognitive Artifacts in Transition: An Analysis of Information Content Changes between Manual and Electronic Patient Tracking Systems. In Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting (accepted). Santa Monica, Calif.: Human Factors and Ergonomics Society.

Rasmussen, J., A.M. Pejtersen, and L.P. Goodstein. 1994. Cognitive Systems Engineering. New York: Wiley and Sons.

Rogers, M., R.L. Cook, R. Bower, M. Molloy, and M.L. Render. 2004. Barriers to implementing wrong site surgery guidelines: a cognitive work analysis. IEEE Transactions on Systems, Man, and Cybernetics: Part A: Systems and Humans 34(6): 757–763.

Sarter, N.B., and D.D. Woods. 2000. Team play with a powerful and independent agent: a full mission simulation study. Human Factors 42(3): 390–402.

Sellen, A.J., and R.H.R. Harper. 2003. The Myth of the Paperless Office. Cambridge, Mass.: The MIT Press.

Sharit, J., L. McCane, D.M. Thevenin, and P. Barach. 2005. Examining Issues in Communicating Patient Care Information across Shifts in a Critical Care Setting. Pp. 1062–1066 in Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Sharp, T.D., and A.J. Helmicki. 1998. The Application of the Ecological Interface Design Approach to Neonatal Intensive Care Medicine. Pp. 350–354 in Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Vicente, K.J. 1999. Cognitive Work Analysis. Mahwah, N.J.: Erlbaum.

Watson, M.O., and P. Sanderson. 2007. Designing for attention with sound: challenges and extensions to ecological interface design. Human Factors 49(2): 331–346.

Wears, R.L., and S. Perry. 2007. Status boards in accident and emergency departments: support for shared cognition. Theoretical Issues in Ergonomics Science 8(5): 371–380.

Wears, R.L., S. Perry, M. Shapiro, C. Beach, P. Croskerry, and R. Behara. 2003. Shift Changes among Emergency Physicians: Best of Times, Worst of Times. Pp. 1420–1423 in Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting. Santa Monica, Calif.: Human Factors and Ergonomics Society.

Wears, R.L., A.M. Bisantz, S. Perry, and R.J. Fairbanks. 2005. Consequences of Technical Change in Cognitive Artefacts for Managing Complex Work. Pp. 317–322 in Human Factors in Organizational Design and Management-VIII, edited by P. Carayon, M. Robertson, B.M. Kleiner, and P.L.T. Hoonakker. Santa Monica, Calif.: IEA Press.

Webster, J.L., and C.G.L. Cao. 2006. Lowering communication barriers in operating room technology. Human Factors 48(4): 747–759.

Xiao, Y. 2005. Artifacts and collaborative work in healthcare: methodological, theoretical, and technological implications of the tangible. Journal of Biomedical Informatics 38(1): 26–33.

# Understanding and Countering the Proliferation of Weapons of Mass Destruction

# Introduction

J. Scott Goldstein
*SAIC*
*McLean, Virginia*

Gregory A. Hebner
*Sandia National Laboratories*
*Albuquerque, New Mexico*

One has only to glance at a newspaper to understand that the proliferation of weapons of mass destruction (WMDs) represents an emerging threat to the safety and security of people around the world. The mere hint that a nation or nonstate organization is trying to acquire a WMD capability, be it nuclear, chemical, or biological, is enough to earn those groups penalties ranging from economic and political embargos to military action. Because of the significant threat to their populations and economic well-being, nations often feel forced to react, even if information on the nature of the threat is not clear or is incomplete. The enormous complexity of the challenge necessitates a multilevel approach that includes public policy and innovative technologies.

The papers in this section address issues associated with understanding the reasons for, and adopting countermeasures to, the proliferation of WMDs. National policies and international diplomacy are both involved in creating an international environment that discourages proliferation and, it is hoped, eliminates the reasons nations may feel the need to acquire WMDs. The authors discuss national and international policy issues to frame the discussion and technical means, such as sensing platforms, that have been developed to provide timely

intelligence on this evolving threat. Although these cutting-edge systems provide a wealth of information that can inform national policies and responses, gaps remain in our capabilities of observing, characterizing, and determining the intentions and motivations of other countries and nonstate organizations.

# U.S. National Security in New Times

Steven D. Nixon
*Formerly of the Office of the Director of National Intelligence*
*Washington, D.C.*

Our national security stands at a critical crossroads. For a generation we confronted the Soviet Union in a Cold War that ended with the collapse of that empire in 1991. Now, 17 years later, religious and ethnic tensions, independence movements, and terrorism have emerged as major security challenges, but our national security establishment has failed to evolve significantly to confront these new challenges. Unlike any other time in history, these challenges are "supercharged" by globalization, the rapid advance and spread of technology around the world, which is empowering small groups and individuals and leading to a new and very dangerous potential for the proliferation of weapons. A looming example is rapid advances in biotechnology and the potential for the relatively easy creation of new and devastating biology-inspired weapons.

To meet the challenges of this age of globalization, we must significantly adapt our national security posture. We cannot win by maintaining our traditional focus on bigger satellites, faster fighter aircraft, or quieter submarines, all of which take decades to deploy at ever-increasing cost. The bottom line is that we must significantly increase our agility, innovation, and collaboration. However, because these qualities run counter to the usual behavior of large government bureaucracies, we should seriously consider the option of creating a variety of small, agile government agencies and outsourcing new activities to small, agile companies.

## WE WON THE COLD WAR, BUT . . .

The end of the Cold War left the United States as sole remaining superpower—at least for the time being. Our economy is the largest in the world, and we enjoy

tremendous competitive advantages in technology and many other areas critical to sustaining a leading position in global affairs.

But without the Soviet Union to confront, our national security apparatus thrashed around throughout the 1990s looking for a new raison d'être. Even though we significantly increased our focus on terrorism after the September 11, 2001, attacks, our security establishment still has essentially the same processes, culture, values, technical capabilities, and organization that were developed to confront the Soviets. And so we are in the classic position of a successful enterprise that is ironically handicapped by its own prior success, a situation that was aptly described in a National Academies (2007) report, *Rising Above the Gathering Storm*:

> There can be no more dangerous place to be than in first place: the one holding that exalted position becomes everyone else's target, and perhaps worse, is the recognized beneficiary of the status quo—and therefore reluctant to promote, or even accept, change.

In this situation, we are extremely vulnerable to disruptive changes in the global security environment. Great companies facing this problem either successfully adapt or go out of business. As Clayton Christianson, famous for his study of disruptive technologies, has described it (Christianson and Raynor, 2003):

> They [industry leaders] pour resources into their core business. They listen to their best customers. And in doing so, industry leaders get blindsided by disruptive innovations—new products, services, or business models that initially target small, seemingly unprofitable customer segments, but eventually evolve to take over the marketplace. This is the innovator's dilemma—and no company or industry is immune.

Even though the current and future security environments demand greater agility from our national security establishment, we have moved in the opposite direction. For example, the first *Corona* optical spy satellite took slightly more than two years from start to first successful launch. Today, it is not at all unusual for a new government satellite program to take more than a decade to achieve first launch. And while we measure our innovation-cycle times in decades, our adversaries, like the insurgents in Iraq, measure theirs in weeks. The *Washington Post* reported on this problem last year (Atkinson, 2007):

> The Improvised Explosive Device struggle has become a test of national agility for a lumbering military-industrial complex fashioned during the Cold War to confront an even more lumbering Soviet system. . . ."If we ever want to kneecap al-Qaeda, just get them to adopt our procurement system. It will bring them to their knees within a week," a former Pentagon official said.

## FROM THE COLD WAR TO GLOBALIZATION: AN INFLECTION POINT FOR NATIONAL SECURITY

Andy Groves, the former CEO of Intel, describes strategic inflection points as points at which a business transitions from the old state of affairs to a new one. These states are often a company's response to what he calls "10x changes" in one or more of the key forces that impact its business.

The world has faced many 10x changes since the end of the Cold War. Nowhere is this more apparent than in the area of technology in which observed trends in accelerating advances are now described as "laws." The most famous of these is undoubtedly Moore's law, which states that the number of transistors that can be placed inexpensively on an integrated circuit doubles every two years. Moore's law has inspired a cottage industry of sorts in producing new technology "laws":

- disk-storage density doubles every 12 months (Kryder's law)
- bandwidth to high-end home users doubles every 21 months (Nielsen's law)
- the amount of data coming out of an optical fiber doubles every nine months (Butter's law)
- the amount of available DNA-sequence data doubles every 18 months (observed, but awaiting someone to attach a name) (Bio Economic Research Associates, 2007)

The United States no longer has a corner on the market in technology. In fact, we are now a net importer of technology products, and these advances are available globally (NSF, 2008). These trends inspired Thomas Friedman (2005) to write the bestseller, *The World Is Flat*, in which he asserts that the diffusion of accelerating technical advances around the globe is creating the ultimate level playing field.

Our friends and adversaries around the world now have access to the same powerful technical capabilities we do. Armed with these new capabilities, small groups and individuals now have the wherewithal to threaten even the mightiest of nations. Consider the impact of the 19 men on September 11 who used modern aviation technology against us. They not only killed thousands of Americans, but also drove our country to spend nearly $1 trillion in response. The amount we have spent on the Iraq and Afghanistan wars already exceeds the amount we spent on Viet Nam, even when adjusting for inflation (Stiglitz and Bilmes, 2008).

Although our national security operations are running at an exceedingly high tempo, I believe the government transformation we need to succeed in the "flat world" is near paralysis. In some sense, this is understandable. Andy Groves (1999) describes this same condition in companies facing strategic inflection points:

> Ideas about the right direction will split people on the same team. After a while, everyone will understand that the stakes are enormously high. There will be a growing ferocity, determination and seriousness surrounding the views the various participants hold. People will dig in. These divergent views will be held equally strongly, almost like religious tenets. In a workplace that used to function collegially and constructively, holy wars will erupt, pitting coworkers against coworkers, long-term friends against long-term friends. Everything senior management is supposed to do—define direction, set strategies, encourage team work, motivate employees—all these things become harder, almost impossible. Everything middle management is supposed to do—implement policy, deal with customers, train employees—also becomes more difficult.

## TRENDS IN BIOTECHNOLOGY: SOON ALMOST ANYONE CAN HAVE A WEAPON OF MASS DESTRUCTION

A good example of the potential threat from globalization is the proliferation of biotechnology that could enable almost anyone with minimal technical savvy to build some pretty scary bio-capabilities. DNA sequencing capabilities are proceeding faster than Moore's law (Bio Economic Research Associates, 2007), and nasty viruses, such as polio and Spanish flu, have not only been sequenced, but have also been artificially reconstructed directly from these sequences (Cello et al., 2002; Taubenberger et al., 2005). The European Molecular Biology Organization reported:

> . . . considering the rapid development of molecular biology, it is only a question of time before the artificial synthesis of agents or new combinations of agents becomes possible. This danger was highlighted last year by a worrying article in *Science*: a research team at the State University of New York in Stony Brook chemically synthesized an artificial polio virus from scratch (Cello et al., 2002). They started with the genetic sequence of the agent, which is available online, ordered small, tailor-made DNA sequences and combined them to reconstruct the complete viral genome. In a final step, the synthesized DNA was brought to life by adding a chemical cocktail that initiated the production of a living, pathogenic virus (Van Aken and Hammond, 2003).

Today, websites such as *www.mrgene.com* offer online DNA synthesis—just submit your sequence to the website, and they will quickly ship your gene. Incidentally, they happen to be running a special in June and July 2008—your gene sequence for just $0.49 per base pair! A headline in *Wired* recently declared, "Scientists Build First Man-Made Genome; Synthetic Life Comes Next" (Madrigal, 2008).

If the prospect of bioterrorism sounds far fetched, consider that there is in fact a long sad history of such attacks. Bioterrorism dates back as far as ancient Rome where dead and rotting animals were thrown into wells to poison water supplies (*http://en.wikipedia.org/wiki/Bioterrorism*). The anthrax

attacks immediately following September 11 are a recent example in the United States; prior to that the Rajneeshee bioterror attack of 1984 sickened 750 individuals in Oregon with salmonella food poisoning (*http://en.wikipedia.org/wiki/1984_Rajneeshee_bioterror_attack*).

Given current advances in technology, it is not difficult to imagine a world a few years from now in which a teenager can create a biological virus almost as easily as today's teenager can create a computer virus. This is indeed a scary future.

## CONCLUSION

How will the national security establishment respond to this threat and others (cyber attack comes to mind) that derive from current technology trends? The advantage now goes to organizations that can operate faster, more innovatively, and more collaboratively than their competitors. In my last job as director of science and technology for the director of national intelligence, we called this "Speed, Surprise, and Synergy." Similarly, Secretary of Defense Robert Gates (2007), in a recent speech, said, "But these new threats also require our government to operate as a whole differently—to act with unity, agility, and creativity."

We must recognize that we no longer live in the industrial age and that traditional industrial-scale solutions simply cannot, in themselves, address many of our current problems. The information age is evolving into a new networking age in which everyone and everything is connected. Our national security establishment must learn to operate across interlinked social networks, financial networks, communication networks, and computer networks. We must make decisions, produce capabilities, and operate at network speed, not industrial speed.

Our national security establishment's love affair with hard science, particularly physics, chemistry, and engineering, must expand to include biology, anthropology, psychology, and other so-called soft sciences. The U.S. government must learn to work more effectively with nongovernment providers and allies. In addition, we should create small, agile, innovative agencies and also outsource more activities to small, agile, innovative companies that do not have to overcome crippling bureaucratic barriers. In short, we must let go of the Cold War way of doing things and move boldly into the 21st century.

## REFERENCES

Atkinson, R. 2007. The Single Most Effective Weapon Against Our Deployed Troops. The Washington Post, September 30, 2007, p. A1.

Bio Economic Research Associates. 2007. Genome Synthesis and Design Futures: Implications for the U.S. Economy. Available online at <http://www.bio-era.net/research/ GenomePurchaseForm.html>.

Cello, J., A.V. Paul, and D E. Wimmer. 2002. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science 297(5583): 1016–1018.

Christianson, C., and M. Raynor. 2003. Innovator's Solution: Creating and Sustaining Successful Growth. Cambridge, Mass.: Harvard Business School Press.

Friedman, T. 2005. The World Is Flat: A Brief History of the Twenty-First Century. New York: Farrar, Straus and Giroux.

Gates, R. 2007. Landon Lecture (Kansas State University). Remarks delivered by Secretary of Defense Robert M. Gates, Manhattan, Kansas, November 26, 2007.

Groves, A. 1999. Only the Paranoid Survive. New York: Doubleday.

Madrigal, A. 2008. Scientists build first man-made genome; synthetic life comes next. Wired, January 24, 2008.

The National Academies. 2007. Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future. Washington, D.C.: The National Academies Press.

NSF (National Science Foundation). 2008. U.S. Trade Balance in High-Technology Goods: 2000–2006. Available online at <http://www.nsf.gov/statistics/seind08/slides.htm>.

Stiglitz, J., and L. Bilmes. 2008. The Three Trillion Dollar War. The New York Times, February 23, 2008. Available online at <http://www.timesonline.co.uk/tol/comment/columnists/guest_contributors/article3419840.ece>.

Taubenberger, J.K., A.H. Reid, R.M. Laurens, R. Wang, G. Jin, and T.G. Fanning. 2005. Characterization of the 1918 influenza virus polymerase genes. Nature 437(7060): 889–893.

Van Akin, J., and E. Hammond. 2003. Genetic engineering and biological weapons. EMBO Reports 4 (Supp1): S57–S60. Available online at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1326447>.

# Combating Weapons of Mass Destruction: Translating Strategic Guidance into Actionable Solutions

CHARLES L. BEAMES
*Space and Intelligence Office*
*Under Secretary of Defense*
*Washington, D.C.*

Soon after the attacks of September 11, 2001, which shattered the sense of domestic security of Americans living in the United States, individual and national security attention was drawn to the question of what else might happen. Quickly, people began to realize that a significant vulnerability of the free and open society we enjoy is the risk from "weapons of mass destruction," or WMDs, in the hands of vengeful terrorists like those who attacked us on 9/11. Along with that realization was a renewed and refocused emphasis on both the nonproliferation and counter-proliferation of WMDs, which are *nothing more than the next turning in a long history of innovation trumping innovation in man's quest to outmaneuver and subdue his enemy.*

## THE IMPORTANCE OF MANEUVER

> Aptitude for maneuver is the supreme skill in a general; it is the most useful and rarest of gifts by which genius is estimated.
>
> – Napoleon Bonaparte

"Maneuver," usually listed second after "Mass" in a list of the Principles of War, is defined in joint doctrine as placing the enemy in the most vulnerable position for the optimal application of force. Today, this 20th-century definition of maneuver must be broadened to include gaining dimensional advantage over an enemy, whether by sea, land, air, space, or cyberspace. Viewing the history of warfare through the prism of technology development and its impact on maneuver brings to light the critical nature of technology in the evolution of warfare.

From the continuous lengthening of standoff range to the introduction of the railroad for logistical resupply, historians have often heralded the successful introduction of a new technology as a turning point in military history, the point at which one side gains an advantage over the other. The stirrup ushered in the age of cavalry, dramatically changing the tempo of war and giving the advantage, at least temporarily, to the land forces that were best able to use it. The same can be said of the internal combustion engine and its application to the tank, which ushered in Blitzkrieg strategies and tactics. Is there any doubt that the successful integration of airpower over the last 100 years, from a largely observational platform with fighter escort in WWI to precision "Shock and Awe" in Operation Iraqi Freedom, has been critical to U.S. strategic dominance in maneuver warfare? The introduction of space-based capabilities in communication, surveillance, and navigation are examples of technologies that have provided significant early warning of enemy positions, movements, and intentions.

In the 21st century, technological development is increasing at unprecedented levels. Unclassified briefings at the highest levels of our national intelligence community indicate that their gravest concern is the combination of technology acceleration (Moore's law in computing power, custom-designed DNA bacteria for the cost of a new car, etc.) and technological leveling through the instantaneous diffusion of information over the Internet and material via overnight shipping.

## INNOVATION, THE HISTORICAL PIVOT POINT IN MANEUVER WARFARE

What becomes clear through a study of maneuver warfare is that more often than not, the most significant and abrupt changes in a combatant's ability to "gain the dimensional advantage through movement" coincides with the successful application of a new technology. Placing an adversary in a disadvantageous posture can be accomplished in one of two ways. First, one can reposition oneself into a position that leaves the opposing military forces in a relatively weaker posture. Second, one can lure the enemy into such a position that he is left with no choice but to move into a disadvantageous position (the classic "horns of a dilemma"). Throughout history, a classic method of the former has been relentless pursuit by military technologists to lengthen standoff range, which enables a force to maintain its strike advantage while making its opponent's force much weaker and less effective.

The original standoff range, when the balled fist was the major weapon, was very likely arm's length. Over the millennia, besides making weapons more lethal, man has sought to lengthen the range of rocks, spears, and so on, culminating in the airplane, the long-range missile, and software viruses, each giving him the ability to outmaneuver his opponent and strike him at will. I say this not to lessen the importance of other factors, such as skill, courage, and reflex, but all else being equal, the first military to adapt a "lengthened knife" or spear to its warfighting apparatus gained a clear advantage over its enemies.

## THE GLOBALIZED THREAT

> The time when only a few states had access to the most dangerous technologies has been over for many years. Dual-use technologies circulate easily in our globalized economy, as do the scientific personnel who design and use them. As a consequence, it is more difficult for us to track efforts to acquire, for nefarious purposes, these components and technologies.
>
> –Annual Threat Assessment of the Director of National Intelligence for the Senate Armed Services Committee, 27 February 2007

According to the director of national intelligence, the biggest threat is technology invention and diffusion and the ability of our adversaries to "live on the exponential technology curve," unlike the U.S. military, which is encumbered by the weight and drag of its bureaucracy and infrastructure. Nimble adversaries can leapfrog a century of science thanks to the diffusion of technological knowledge and access to weapons and materials largely as a result of the Internet and the explosive growth of the Google and eBay economies. Ominously, unanticipated changes in the character of war can be major pivot points in political history as well.

Richard Hellie, in his book *Warfare, Changing Military Technology, and the Evolution of Muscovite Society,* describes how migrating from the bow and arrow to the musket not only led to victory in warfare, but also resulted in a reordering of political power. From 1450 to 1725, Russia experienced two revolutions as a direct result of military threats made possible largely by the full-scale introduction of technological advances. The first threat was from the Tatar light cavalry, and the second was from the Swedish infantry. The replacement of the light bow and arrow with the musket changed the nature of warfare from resource control to territorial control. Russia quickly reordered its society into a very rigid, castelike system and was able to defend itself against threats from Lithuania, Poland, and Sweden.

We live in an epoch unparalleled in human history, when "virtually anything of value is offered in today's global marketplace—including illegal drugs . . . machine guns and rocket launchers, and centrifuges and precursor chemicals

used in nuclear weapons development" (Naim, 2005). The shift to a global arms market means a "massive transfer of goods and equipment once under the exclusive control of national armies into private hands released into the market products ranging from rockets launchers to SCUD missiles and nuclear designs and machinery" (Naim, 2005).

## MANEUVER, ADVANTAGE, INNOVATION, AND UNCERTAINTY

> It's tough to make predictions, especially about the future.
>
> – Yogi Berra

Clearly, in warfare, as in the commercial marketplace, the inventor of a technology doesn't always "win." The winner is the successful innovator. Whether in standoff range or logistics, reconnaissance or precise navigation, the military that most successfully applies the invention to the art of war gains the advantage and wins, all other instruments of political power being equal.

## INNOVATION IN THE COMMERCIAL MARKETPLACE

As part of a thesis for the National War College, this author conducted comparative case studies in innovation in three industry leaders, Google, Apple, and IBM, to determine their common cultural characteristics that could be applied to the defense world to improve the nation's ability to innovate solutions to counter the asymmetric technologies being used to significantly degrade U.S. military power. The studies were conducted using the framework devised by Jim Collins in his treatise, *Good to Great*, and the results are summarized below.

### The 11 Cultural Lessons of Innovation

#### People

- Recruit the best people you can possibly afford, and avoid the worst at all costs.
- Establish a culture of recognizing the problems to be solved.
- Direct a sizable and conscious amount of work time and effort to fostering creativity.
- Organize around small work groups of three or four people to encourage the highest level of innovation.
- The most important leadership trait is credibility; a leader should be technically skilled and perceived as such.

### Thinking

- A commitment should be made at all levels to "radical honesty."
- The starting point of all relevant innovation is solving a real problem.
- The essence of creative, innovative thinking is accomplished by individuals.

### Actions

- Ensure that there is disciplined, rational, but audacious movement at all levels.
- Vision and *raison d'être* should be internalized by all members of the organization at every level.
- Everyone must have hands-on skills related to the work they oversee or are involved with.

## NONPROLIFERATION AND COUNTERPROLIFERATION OF WEAPONS OF MASS DESTRUCTION

Few propositions are more fraught with disaster than the suggestion that the military be run like a business. The goals, motivations, necessity, and reasons for existence of companies and the military often run in opposite directions. Nevertheless, the confluence of history at which this nation finds itself is unique. Scientific and technological know-how and invention are at their highest rate in the history of the world, and, at the same time, the world is becoming a technologically level playing field. This "flattening" leads to a rise in productivity and living standards for people around the world, but it is also one of the factors that leads to technological advantage. For that reason, the national security apparatus responsible for combating WMDs ought to consider these factors as it develops and devises organizations for to control and counter the spread of WMDs.

The Internet is the engine that drives leveling and advancement. Although most technologists argue that the Web-enabled world is still in its infancy, sufficient time has elapsed that a highly inventive and competitive economy has formed around it. Because of the lower capital costs associated with entering this market, it is likely to remain highly competitive. The intellectual property associated with it is predominantly in software, and the economic rewards can be asymmetric. As a result, the Web-based world is a highly dynamic, competitive, and innovative environment.

The cultural attributes that characterize innovative organizations also apply to terrorist organizations trying to obtain WMDs. Thus, we must call into question some general assumptions of the culture on which our highly bureaucratic national security establishment is based. Even though this culture, organization, and tradition is in some ways contrary to conventional wisdom, our nation is dealing with

a new distribution of power based on technological leveling, and we ignore this shift at our peril. We need further studies of this situation, but they must be done soon. As Napoleon said, "Given the same amount of intelligence, timidity will do a thousand times more damage in war than audacity."

Philip Howard says in his book, *The Death of Common Sense*, that very little gets done in mature bureaucracies because processes put in place over the years have stripped responsibility from bureaucrats leaving them unaccountable and not apt to display the three attributes most necessary to solving problems and getting things done: effort, courage, and leadership. Without rethinking the reward structure of our military and beginning to move away from a highly bureaucratized, static organization toward a flattened, empowered, versatile, and highly innovative culture, there is significant risk that we will be caught by surprise by grave threats to our national survival, much as our army was caught by surprise by the evolution of improvised explosive devices in Iraq.

## CONCLUSION

The only limiting factor for military application of technology is the imagination of a certain percentage of the 6.5 billion people who would rejoice in seeing the United States humiliated for its perceived hubristic behavior. The question that ought to be haunting the nonproliferation and counterproliferation establishment is how we can be ready for a 21st-century "Mongol" (the one who applied the stirrup to cavalry warfare and changed the world), who at this moment may recognize the military utility of something that no one else sees and that can bring down the most powerful nation the world has ever known.

## SELECTED BIBLIOGRAPHY

### Books

Angevine, R.G. 2004. The Railroad and the State: War, Politics, and Technology in 19th Century America. Stanford, Calif.: Stanford University Press.

Berry, F.C., Jr. 1993. Inventing the Future: How Science and Technology Transform Our World. New York: Macmillan.

Biddle, S.D. 2004. Military Power: Explaining Victory and Defeat in Modern Battle. Princeton, N.J.: Princeton University Press.

Clausewitz, C. von. 1976. On War. Translated by Michael Howard and Peter Paret. Princeton, N.J.: Princeton University Press.

Collins, J.C. 2001. Good to Great. New York: Harper Collins.

Greene, R. 2006. The 33 Strategies of War. New York: Penguin Group.

Hellie, Richard. 1990. Warfare, Changing Military Technology, and the Evolution of Muscovite Society. Pp. 74-99 in Tools of War, Instruments, Ideas, and the Institutions of Warfare, 1445-1871, edited by John A. Lynn. Urbana, Ill.: University of Illinois Press.

Howard, P.K. 1994. The Death of Common Sense. New York: Random House.

Leonhard, R.R. 1998. The Principles of War for the Information Age. Novato, Calif.: Presidio Press.

Leonhard, R.R. 1994. Fighting by Minutes: Time and the Art of War. Westport, Conn.: Praeger.

Lieber, K.A. 2005. War and the Engineers: The Primacy of Politics over Technology. Ithaca, N.Y.: Cornell University Press.

Lynn, J.A. 1990. Tools of War: Instruments, Ideas, and Institutions of Warfare, 1445–1871. Urbana, Ill.: University of Illinois Press.

Naim, M. 2005. Illicit: How Smugglers, Traffickers, and Copycats are Hijacking the Global Economy. New York: Doubleday.

Porter, M.E. 1990. The Competitive Advantage of Nations. New York: Free Press.

Sun Tzu. 1963. The Art of War, translated by Samuel B. Griffith. New York: Oxford University Press.

Weigley, R.F. 1973. The American Way of War. Bloomington, Ind.: Indiana University Press.

## Government Publications

Joint Chiefs of Staff. 2000. Joint Publication 1: Joint Warfare of the Armed Forces of the United States. 14 November 2000. Washington, D.C.: Joint Chiefs of Staff.

Joint Chiefs of Staff. 2006. Joint Publication 3-0, Joint Operations. 17 September 2006. Washington, D.C.: Ft. Belvoir Defense Technical Information Center.

Nixon, S. 2006. Living in a World of Rapid Change: The IC S&T Transformation Imperative. Briefing to Congress, August 18, 2006.

Office of the Director of National Intelligence. 2007. Annual Threat Assessment of the Director of National Intelligence for the Senate Armed Services Committee, February 27, 2007.

## Online Sources

McGregor, J., M. Arndt, R. Berner, I. Rowley, K. Hall, G. Edmondson, S. Hamm, M. Ihlwan, and A. Reinhardt. 2006. Innovation: The World's Most Innovative Companies. <http://www.businessweek.com/magazine/content/06_17/b3981401.htm>.

Negroponte, N. 2003. Creating a Culture of Ideas. <http://www.technologyreview.com/Biztech/13074/>.

The Quotations Page. Yogi Berra. <http://www.quotationspage.com/quotes/Yogi_Berra/>.

# Nuclear Deterrence in the 21st Century: The Role of Science and Engineering

Joseph C. Martz and Jonathan S. Ventura
*Los Alamos National Laboratory*
*Albuquerque, New Mexico*

For five decades after the Second World War, the role of nuclear weapons in the U.S. defense posture was largely to deter the Soviet Union from attacking the United States, its allies, and friends. Almost as important was the nuclear umbrella (security assurances) the United States provided to discourage (dissuade) non-nuclear nations from pursuing nuclear weapons. These two roles were critical to the formulation of the Nonproliferation Treaty.

When the Cold War ended with the sudden collapse of the Soviet Union, the purpose of nuclear deterrence was fundamentally changed. Although the United States continued to retain nuclear weapons as instruments of retaliation, the dynamics of deterrence were evolving. In the future, we will need substantially fewer deployed weapons to address a variety of potential adversaries.

Unlike the security challenges of 50 years ago, 21st-century challenges are multipolar and often asymmetric. However, a few key countries still have substantial nuclear arsenals and active, growing nuclear weapons programs that threaten vital U.S. interests. Thus maintaining a credible U.S. nuclear deterrent in the eyes of our allies and our adversaries, as well as supporting the goal of nonproliferation, will continue to be critical to U.S. national security.

Overlaid on this national security backdrop are concerns about the effects of

a growing world population on climate, food, and energy—which greatly complicate the security challenges facing policy makers today. This changing paradigm requires that the United States take additional measures (1) to assure our allies that we are still a trusted security partner and (2) to deter potential adversaries from taking aggressive action that could threaten global stability. Every U.S. president since Truman has affirmed the role of nuclear weapons as a supreme deterrent and protector of last resort of U.S. national security interests.

Recently, President Bush called for a "…credible deterrent with the lowest possible number of nuclear weapons consistent with our national security needs, including our obligations to our allies."[1] How can this be achieved? Can we continue on a path of nuclear reductions while maintaining the national security benefits of nuclear deterrence? Science and engineering will play a key role in the new paradigm for nuclear deterrence, "capability-based deterrence."

## CAPABILITY-BASED DETERRENCE

Capability-based deterrence is based on the principle that an agile, repeatedly demonstrated capability to develop and produce deployable nuclear weapons will greatly strengthen the deterrent and enable meaningful reductions in the size of the total stockpile. In this scenario, the country can rely, in part, on a working weapons complex that could deliver limited numbers of nuclear weapons should the situation require rather than on large numbers of reserve or deployed warheads for contingency purposes.

The new strategy would provide the benefits of deterrence while enabling us to meet some key goals, such as reducing the stockpile of nuclear weapons. Indeed, the head of the National Nuclear Security Administration (NNSA) stated in December 2007—when he proposed a transformation of the nuclear weapons complex—that "the United States' future deterrent cannot be based on the old Cold War model of the number of weapons. Rather, it must be based on the capability to respond to any national security situation, and make weapons only if necessary."[2] This adoption of capability-based deterrence would represent a shift in the emphasis of our nuclear policy. The role of science and engineering would become a critical element in establishing the agility and confidence necessary for this strategy to work.

The principal elements of capability-based deterrence are (1) the weapons themselves (albeit fewer and potentially designed to meet the specific requirements of this strategy); (2) the design, development, and manufacturing elements of the weapons complex. It is not only the capabilities of our military forces that

[1] *Remarks by the President to Students and Faculty at National Defense University*, Fort Lesley J. McNair, Washington, D.C., May 1, 2001.

[2] *Remarks by Thomas D'Agastino on the Introduction of the Complex Transformation PEIS*, U.S. Department of Energy Headquarters, Washington, D.C., December 17, 2007.

assure our allies and deter potential adversaries; it is also our capability to sustain and modernize our forces, while demonstrating the ability to respond rapidly to new or emerging threats.

The notion of capability-based deterrence is not completely new. It was included in the Bush administration's "new strategic triad" concept and was emphasized in the administration's 2001 Nuclear Posture Review (NPR). It was also part of the Clinton administration's NPR in 1994 and was a principle in the founding of the Stockpile Stewardship Program.

This science-based program of experimentation, improved diagnostics, and greatly increased computational capabilities gave us the tools to assess and redress problems in our management of the stockpile. Research conducted at new experimental facilities has improved our understanding of the physics of weapons performance and thus contributed to the program's remarkable technical success, which has increased confidence in our ability to transition to capability-based deterrence.

## ENABLING A CAPABILITY-BASED DETERRENT

Timeliness and agility are critical elements of this new strategy. We must be able to detect and respond to a potential adversary quickly enough to counter any provocative act. The need for timeliness is an interesting contrast to our Cold War posture. In decades past, we had to be ready to respond to a provocative act on a moment's notice. Thus we had bombers on constant standby, intercontinental missiles on hair-trigger alert, and submarines on continuous patrol in the great oceans. Our answer to the timing question then was in minutes.

Today, a threat that might require a response of such a large arsenal may not become manifest for several years. Indeed, in an environment of stockpile reductions—both the United States and Russia have reduced their arsenals by more than 90 percent from their Cold War peaks—we no longer need the large nuclear forces that characterized the Cold War. If an adversary decided to restart an arms race, it would require a large investment on their part, and, in principle, sufficient time for the United States and its allies to respond.

In essence, the change in strategy would mean moving from the deterrence afforded by large numbers of deployed and reserve weapons to deterrence provided by a smaller number of deployed weapons and a robust and agile infrastructure and capability. This strategy could potentially provide many of the benefits of nuclear deterrence while enabling us to continue to reduce our stockpile.

The promise of this strategy—the ability to provide an agile, diverse response to many threats—would provide us with an advantage we did not have when we relied solely on a stockpile of Cold War-optimized, high yield-to-weight weapons. Science and engineering will be key to enabling the United States to become an agile responder, because a capability-based deterrent must be grounded in science.

The challenge is to develop and demonstrate a steady-state capability to execute a complete cycle of warhead design, certification, development, and production in a three-to-seven-year time frame. This cannot be done with our current very outdated and archaic production complex and 1970s-era design practices.

The NNSA proposal for a transformation of the nuclear weapons complex that will increase U.S. agility and inspire more confidence is not entirely theoretical. Recent developments in the NNSA complex have demonstrated the viability of some key elements of this strategy, most notably a few recent stockpile life-extension program activities, as well as the reliable replacement warhead (RRW) feasibility study.

## THE RELIABLE REPLACEMENT WARHEAD

In the RRW study, laboratory design teams were able to provide highly mature designs in less than 12 months—largely thanks to modern engineering and design tools created under the Stockpile Stewardship Program. In addition, the laboratories built prototype demonstration hardware and conducted limited, nonnuclear proof tests of their designs within 18 months. Note that in the 1980s, this level of design maturity required several years. These proof tests exercised a portion of the production complex and provided a concrete example of the agility and timeliness that would be possible.

In addition, RRW designs were based on the "relaxation" of a Cold War objective—maximizing yield-to-weight ratios. The RRW designs instead increased performance margins and backed away from known failure modes of the legacy stockpile. Increased performance margins coupled with advances in weapons science will provide high-confidence in certification of these RRW designs without nuclear testing. In the future, advanced features could be included in the designs to improve safety, security, and use control in the warhead and also improve the efficiency of manufacturing operations.

## SUMMARY

A movement toward an active, fully functional, demonstrated capability-based deterrence program could help the nation meet its future policy objectives: further reductions in the size of the stockpile, certification of our nuclear deterrent without nuclear testing, and advancing compliance with Article 6 commitments under the Nuclear Nonproliferation Treaty.

Two key enablers of an articulated capability-based deterrent are the transformation of the weapons complex (as proposed by NNSA) and the adoption of many of the concepts and approaches demonstrated in the RRW project—all of which were made possible by the science and engineering methods and tools developed under the Stockpile Stewardship Program.

Although the U.S. government has proposed that we move toward a "…credible

deterrent with the lowest-possible number of nuclear weapons consistent with our national security needs," we can only move in that direction in an environment in which our security is maintained, our allies are assured of our commitments, and our adversaries are dissuaded and deterred. However near or remote a world free of nuclear weapons may be, a capability-based deterrence to meet today's threats can facilitate reductions in the stockpile while maintaining our security and limiting technical risks.

## RECOMMENDED READING

Shultz, G.P., W.J. Perry, H.A. Kissinger, and S. Nun. 2008. Toward a Nuclear Free World. Editorial in the *Wall Street Journal*, January 15, 2008.

Lewis, J. 2008. Minimum deterrence. Bulletin of the Atomic Scientists 64 (3): 38–41.

Bunn, M.E. 2007. Can deterrence be tailored? Strategic Forum 225: 1–8.

# Dinner Speech

# Energy Policy and the Role of Technology in National Security[1]

A.D. Romig, Jr. with Arnold B. Baker[2]
*Sandia National Laboratories*
*Albuquerque, New Mexico*

Global security is a dynamic, complex system of systems. From one perspective, it is a global/national political-economic-technology system. From another, it is a defense, intelligence, and security system. From yet another perspective, it is an energy (coal, oil, natural gas, and renewables) and environmental system that interacts with our economies, demographics, technology, and national and geopolitical systems. Consequently, U.S. energy security requires thoughtful analysis and global engagement in all of these interrelated dimensions, rather than just the "energy" or "energy-environmental" dimensions. As a practical matter, globalization, growing international interdependencies, and geopolitics set the context for meaningful discussions about energy security.

[1]Sandia National Laboratories is a multiprogram organization operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy National Nuclear Security Administration, under contract DE-AC04-94AL85000.

[2]Dr. Romig is executive vice president, deputy laboratories director for Integrated Technologies and Systems, and interim chief operating officer, Sandia National Laboratories. Dr. Baker is chief economist, Sandia National Laboratories. This is an extended abstract of the keynote address by Dr. Romig to the National Academy of Engineering U.S. Frontiers of Engineering Symposium, University of New Mexico, September 18, 2008.

For example, the economic balance of power in today's global economy appears to be shifting from current "developed" economies (members of the Organisation for Economic Co-operation and Development, or OECD) to "developing" economies, such as China and India. OECD members currently account for more than half the world's GDP (on the basis of purchasing power parity, or PPP), up from only 30 percent in 1700. When current economic trends are extended to 2050, OECD members' share of world GDP could very well decline to approximately 30 percent, while current non-OECD members would account for 70 percent, as they did in 1700. At the same time, the U.S. share of world GDP in 2050 could decline from the current 20 percent (PPP basis) to 10 percent, while China's share could grow to 30 percent, up from about 10 percent today.

This economic power shift is being driven by growing global economic integration and interdependency, despite continuing protectionist threats. In the future, nations will be both more competitive and more cooperative. In a more competitive world, the scope of national policies with major economic impact may become increasingly limited, while the need for clear domestic consumer-producer energy price signals and consistent energy-security, environmental, and economic objectives and policies will become more important.

Driven in part by the Internet and economic integration, the world has become increasingly complex geopolitically (Figure 1). While the U.S. remains

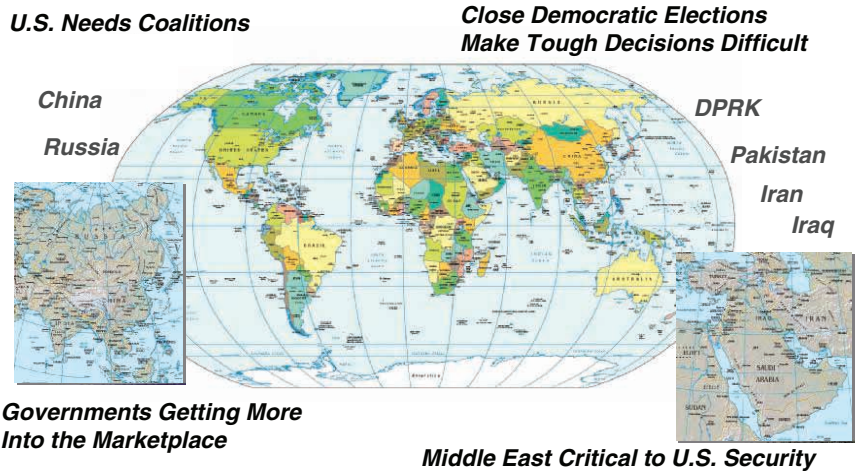


FIGURE 1 Map showing the increasing complexity of geopolitics. Source: Sandia National Laboratories.

the world's only superpower, it increasingly needs coalitions to support its global security efforts, such as in the Middle East, Afghanistan, Iran, and North Korea, in dealing with emerging Russian adventurism, and in domestic and international financial markets. All of these engagements are critical to U.S. security. And, while close elections and partisan politics make forging a national consensus on complex issues such as energy security and climate change difficult in democratic countries, autocratic governments such as in Russia and Venezuela have returned to the energy business and are aggressively using energy as a tool of national policy.

At the same time, the character and nature of global conflicts have changed, and they will continue to evolve. For example, in the current geopolitical environment physical force alone may not resolve conflicts, and philosophical differences may be irreconcilable. Terrorism has become a reality, and an "unwillingness to kill" is viewed by some as weakness. Yet geopolitical and national security systems remain flexible to minimize conflict and permit economic development to progress.

Between 2005 and 2030 (Figure 2), world energy demand and carbon emissions are expected to grow more than 50 percent, and consistent with the economic trends previously mentioned, developing countries will account for 80 percent of that increase according to the Energy Information Office. In 2030, liquid fuels, mostly petroleum based, would account for 33 percent of world energy demand, down slightly from 36 percent today. The share of coal would grow from 27 percent today to 29 percent, while the share of natural gas would remain at approximately 23 percent. Since renewables (including hydroelectricity) and nuclear power would remain at 8 and 6 percent respectively, the share of global energy from fossil fuels would remain at 86 percent.

From an energy-security perspective, it is important to note that today's fossil-energy reserves are geographically concentrated. Some 60 percent of proven oil reserves are in the Persian Gulf and Russia—not the most stable regions in the world. Saudi Arabia, Iran, and Iraq account for 20, 10, and 9 percent, respectively, while Russia accounts for 5 percent. This same region also accounts for 68 percent of proven natural gas reserves. Russia has 27 percent, while Iran and Qatar have 16 and 15 percent, respectively. Hence disturbances in the Persian Gulf and Russia will not only affect world oil markets, but will also affect the world natural-gas markets, as the international market for liquefied natural gas grows.

Coal is more widely dispersed geographically. The United States has 27 percent of proven reserves, while Russia, China, and India have 17, 13, and 10 percent each. However, because burning one BTU of coal releases 80 percent more carbon dioxide than a BTU of natural gas, and about 40 percent more than a BTU of oil, increasing restrictions on coal use are likely to be imposed in some countries.

Fortunately, according to the International Energy Agency (IEA), while the world produced some trillion barrels of oil by 2005, there are about 4.5 trillion barrels of oil yet to be produced. This includes OPEC Middle East oil, other con-

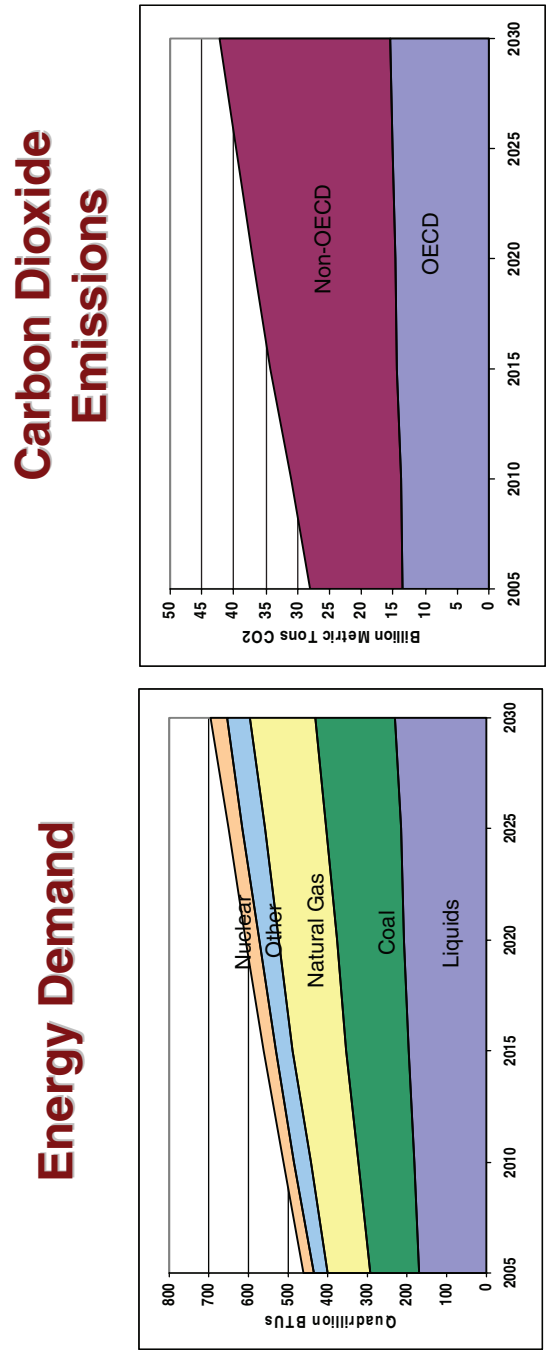


FIGURE 2 Graphs showing estimated increases in world demand for energy and in carbon emissions by 2030. Source: U.S. DOE EIA (2008).

ventional oil, heavy oil and bitumen, oil shale, oil in the Arctic, and increases in oil from enhanced recovery techniques.[3] Most of this oil can be produced for less than $50 to $70 per barrel ($2004). In addition, and not counted in these totals, are liquid fuels that can be produced from coal, natural gas, and biological materials. However, it will take time, some improved technologies, and sizable capital investments to bring these various liquids on stream.

The same IEA study indicated that known resources of natural gas will last for many years. While 80 trillion cubic meters (TCM) have been produced, 370 TCM, or more than 120 years supply at current consumption rates (2.9 TCM/year) remain. In addition, there are at least 248 TCM of nonconventional gas from coalbed methane, tight gas, and gas shales remaining. (Reliable worldwide estimates for nonconventional gas are not available, so these resources could be two to three times larger.) In addition, between 1,000 and 10,000,000 TCM of gas are locked in the form of hydrates in the seabed and permafrost, but their recovery status is unknown.

Simply having reserves and the ability to extract them, however, does not guarantee a secure supply that will not be disrupted. Physical protection of energy infrastructure—pipelines, tankers, and electricity—presents some unique security challenges, because infrastructure components are widespread, highly visible, and accessible. Many transportation and delivery nodes and links are exposed and in unstable and/or unfriendly regions. In addition, growing energy markets and integration will stretch infrastructure systems and add complexity to their operation and security. Nonetheless, we are making progress. We already have a wide range of tools, with more in development, to help protect energy infrastructure.

Research and development programs are under way to make advanced biofuels from algae and cellulosic ethanol cost competitive and to explore "sunshine to petrol" (Figure 3), an advanced concept that would use solar-energy-powered catalytic reactors with water and carbon dioxide to make synthetic gasoline. Such technology would be key to minimizing our carbon footprint.

Although energy security is a challenging problem, the policy on global climate change is even more of a problem. To stabilize the atmospheric concentration of carbon and other greenhouse gases at current levels, which would ensure that human influence on climate would get no worse than it is today, would require a 50 to 90 percent reduction in current emission levels, according to the Intergovernmental Panel on Climate Change.[4] In effect, without carbon sequestration, the world would have to reduce its current use of fossil fuels by 50 percent or more; and unless developing countries like China also reduce their current use of fossil fuels by this amount, then the U.S. and the rest of the world would have to make even greater reductions.

[3]IEA, *Resources to Reserves—Oil and Gas Technologies for the Energy Markets of the Future*, 2005.

[4]IPCC, *Climate Change 2007 Synthesis Report,* 2007.

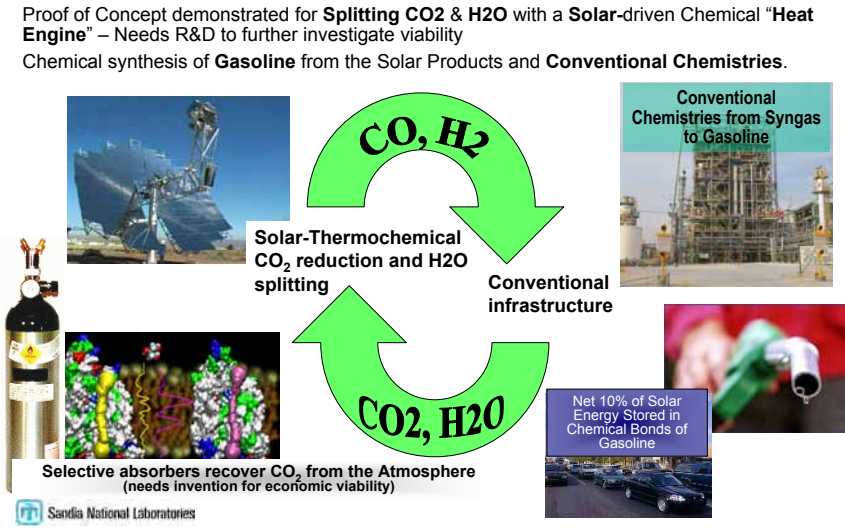


FIGURE 3 Work is being done on producing a carbon-neutral, renewable gasoline. Source: Sandia National Laboratories.

Nuclear energy, through an integrated nuclear-power enterprise, can play a significant role in both energy security and in reducing carbon emissions. Sandia is helping to bring this about through efforts to ensure the safety and security of nuclear facilities, to solve the nuclear waste problem, to provide innovative nuclear-power options, and to help prevent nuclear proliferation (Figure 4). The latter issue will become increasingly important as nuclear power grows worldwide. More enlightened ways of managing the nuclear fuel cycle and nuclear waste will be key to minimizing our nuclear footprint.

Although our current near-term options for dealing with energy (especially oil and natural gas) security and carbon emissions are limited, we believe that a range of technology innovations will ultimately enable significant advances in energy security and reductions in carbon emissions (Figure 5). These advances will both enhance energy supply and reduce energy consumption, help improve the security of our energy infrastructure, and help reduce the carbon footprint. These advances will include high-performance computing, advanced robotics, advanced modeling and simulation, and microelectronic systems.

Some governments and automobile companies are working toward a future hydrogen economy, which they believe will address many of the current energy-security and environmental issues related to our reliance on fossil fuels. A hydrogen economy could improve a number of current problems by reducing our oil dependence on the Middle East and Russia, reducing our fossil-based carbon

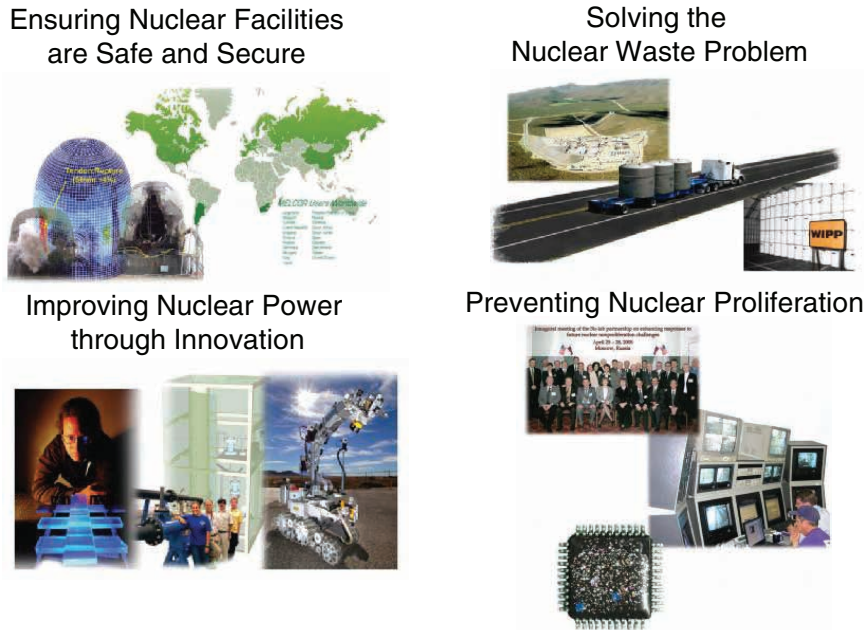


FIGURE 4 Illustration showing the components of an integrated nuclear-power enterprise. Source: Sandia National Laboratories.

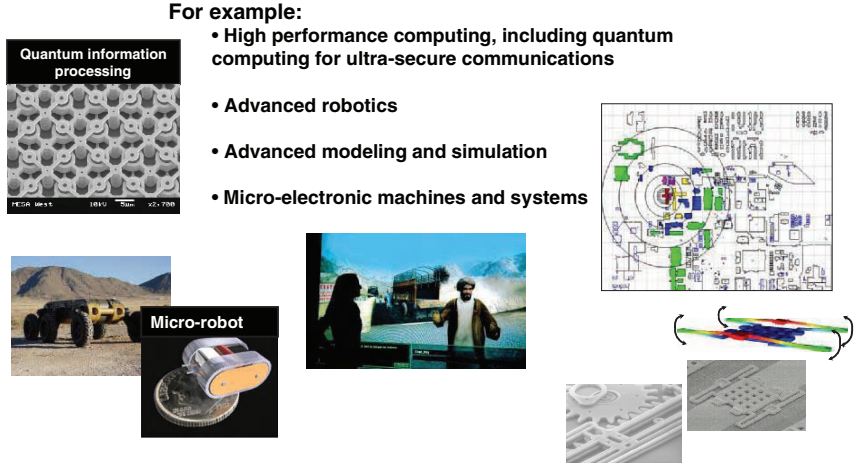


FIGURE 5 Technology innovations for a secure energy future. Source: Sandia National Laboratories.

emissions, and helping to overcome "fixed resource" limitations of fossil fuels and their uneven distribution among countries. However, many hurdles would have to be overcome, such as onboard hydrogen storage, the lifetime of fuel cells, hydrogen production economics, lack of a hydrogen infrastructure, and carbon capture and storage, if the hydrogen is derived from fossil fuels.

As a result, hydrogen fuel and vehicle systems are unlikely to have meaningful market penetration until at least the mid-2020s. In the meantime, automobile manufacturers and others are seriously interested in developing and marketing a wide range of plug-in hybrid and battery-powered electric vehicles that may substantially decrease the use of petroleum-based fuels for automobiles. Although only a limited number of models is available today, most auto companies plan to offer additional models and options by 2010.

From a longer-range perspective, several advanced technologies are on the horizon that could very positively affect our energy system. Nanotechnology, in particular, has the potential to change energy supply and demand in ways we have only begun to consider. For example, solid-state lighting using quantum dots could cut power use for lighting by half. Ultra-high-strength, lightweight nanophase materials could improve car and airplane efficiency substantially. Nanoparticles and nanoarchitectures for energy conversion and storage may offer solutions to low-cost fuel cells and batteries.

## SUMMARY AND CONCLUSIONS

The world economy and energy markets will become increasingly integrated and interdependent, although clearly the risk of "pull-back" and protectionism remains. Based on current trends, energy use and carbon emissions will increase substantially, driven by the developing world. In the near to medium term, the potential for supply shocks and price instability in oil and natural gas will increase. In addition, nuclear power will grow, and nuclear technology will spread, increasing the risk of proliferation. Defense and military complexity will also increase, as will requirements for sound, timely intelligence. However, at the same time, major new energy-technology platforms based on renewables could transform economies and lead to the emergence of other energy markets.

As both economic competition and cooperation intensify, the appetite for high-cost public policies in the United States that are inconsistent with competitor countries will likely become more limited. At the same time, the need for consistent energy-security, environmental, and economic objectives and policies will grow.

The protection of energy infrastructure will continue to be a critical component of national security, and tools are being developed and improved to help provide that protection. Systems analysis, enhanced intelligence, and, as a last resort, military force may be brought to bear; and new technologies will enable new creative solutions to enhance protection even further.

During the transition to more advanced energy and environmental technologies, international flexibility, cooperation, and partnering in many areas, including defense, intelligence, nonproliferation, public policy, and science and technology investment, will be critical to avoiding disruptions in energy supplies. Flexibility, cooperation, and partnering will also be necessary to support international economic and political security, improve the health and well-being of the developing world, and provide a foundation for global and regional economic prosperity and environmental sustainability.

## READING LIST

British Petroleum. 2007. BP Statistical Review of World Energy 2007. June 2007. London, U.K.

EIA (Energy Information Administration). 2007. International Energy Annual. June 21, 2007. Washington, D.C.: U.S. Department of Energy.

EIA. 2008. International Energy Outlook 2008. Washington, D.C.: U.S. Department of Energy.

IEA (International Energy Agency). 2005. Resources to Reserves: Oil and Gas Technologies for the Energy Markets of the Future. Paris, France: OECD (Organization for Economic Cooperation and Development).

IPCC (Intergovernmental Panel on Climate Change). 2007. Climate Change 2007 Synthesis Report. November, 2007. Geneva, Switzerland: IPCC.

Maddison, A. 2001. The World Economy: A Millennial Perspective. Paris, France: OECD.

# APPENDIXES

# Contributors

**Charles L. Beames,** a colonel in the U.S. Air Force, is currently director, Intelligence, Surveillance, and Reconnaissance (ISR) Division, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, where he is responsible for overseeing the acquisition of all space-based ISR programs for the U.S. Department of Defense and the intelligence community. As a design engineer and project manager, he has experience in the development, deployment, and operation of airborne, ground-based, and space-based systems. Col. Beames has been sent to locations around the world in support of ongoing operations. Previously, he was chief, Operational Capabilities Division, U.S. Strategic Command Center for Combating Weapons of Mass Destruction (WMD), which is responsible for identifying and advocating ways of combating WMD operational capabilities and translating strategic guidance into real-world solutions. Earlier in his career, he had a fellowship at the National War College, was the space and intelligence liaison to Congress for Air Force appropriations and held staff assignments at HQ, Air Force Materiel Command, U.S. House of Representatives and the Office of Secretary of the Air Force. He graduated from the U.S. Air Force Academy with a B.S. in engineering mechanics and materials science (1988) and received an M.S.

in mechanical engineering from Johns Hopkins University and an M.S. in national security strategy from the National War College in Washington, D.C.

**Ann M. Bisantz** is an associate professor in the Department of Industrial and Systems Engineering, University at Buffalo, State University of New York (SUNY-Buffalo). Her research interests include the design of human-computer systems using visual and multimodal displays, methods in cognitive engineering, and modeling of dynamic decision making to communicate uncertainties, especially in complex environments, such as military systems and health care settings. Dr. Bisantz is the recipient of a CAREER Award from the National Science Foundation and a Young Investigator Award from SUNY-Buffalo in 2002. She is a scientific editor of *Applied Ergonomics*; program chair for the Cognitive Engineering and Decision-Making Technical Group of the Human Factors and Ergonomics Society; director of undergraduate studies for the Industrial and Systems Engineering Department; co-editor of *Applications of Cognitive Work Analysis* (Taylor and Francis, 2008); and serves on the editorial boards of *Human Factors*, *International Journal of Industrial Ergonomics*, and *Cognitive Engineering and Decision Making*. Dr. Bisantz received an M.S. from SUNY-Buffalo and a Ph.D. from the Georgia Institute of Technology, both in industrial engineering.

**Ronald Laurids Boring,** a senior member of the technical staff in the Risk and Reliability Analysis Department at Sandia National Laboratories, was previously a human-factors scientist working in the Human Factors, Instrumentation and Control Systems Department, Idaho National Laboratory, where he specialized in using human-reliability analysis and cognitive modeling to reduce human error in complex technological environments. His previous experience includes program manager for human reliability research projects funded by the U.S. Nuclear Regulatory Commission and the National Aeronautics and Space Administration and visiting human-factors scientist at the Halden Reactor Project in Halden, Norway. After completing a bachelor's degree in psychology and German at the University of Montana, Dr. Boring was a Fulbright Scholar at the University of Heidelberg. He later completed his masters-level studies in experimental psychology at New Mexico State University and his Ph.D. at the Carleton University Institute of Cognitive Science, thus bringing together his interests in computer science and psychology. He is the author of more than 70 technical articles in the fields of human factors and human reliability.

**Xiaohu Gao,** an assistant professor of bioengineering in the Center for Nanotechnology and Department of Bioengineering, University of Washington (UW), conducts research in molecular bioengineering, nanotechnology, medical imaging, and image-guided therapy. He joined the faculty at UW in 2005 after completing postdoctoral fellowships in the Department of Biomedical Engineering, Georgia Institute of Technology, and the Winship Cancer Institute, Emory University. He

received a CAREER Award from the National Science Foundation in 2007 and the New Investigator Award from the Congressionally Directed Medical Research Programs and Wallace H. Coulter Translational Research Award, both in 2006. He is a member of the American Chemical Society, Biomedical Engineering Society, and American Society of Mechanical Engineering and a reviewer for several journals. Dr. Gao received his Ph.D. in bioanalytical chemistry from Indiana University, Bloomington, in 2004.

**Stephanie Guerlain** is an associate professor in the Department of Systems and Information Engineering at the University of Virginia, where she specializes in the design of decision-support systems, cognitive-systems engineering, human-computer interaction, and data visualization. Her research results have been used in medical, military, process-control, and bioinformatics settings. She has published more than 40 technical papers on various aspects of cognitive-systems engineering, focused on decision-support systems, data visualization, and computer-based training. Prior to joining the faculty at Virginia in 1999, Dr. Guerlain was principal research scientist at Honeywell Technology Center, where she worked primarily on process-control applications. She is a member of the IEEE Systems, Man and Cybernetics Society; Cognitive Science Society; and Human Factors and Ergonomics Society. She and her students have won five conference "best paper" awards and one journal "best paper" award. She received her M.S. and Ph.D. in industrial and systems engineering from Ohio State University.

**Jeff Hrkach,** now vice president of pharmaceutical sciences at BIND Biosciences in Cambridge, Massachusetts, was previously senior director of drug delivery and strategic product development at Momenta Pharmaceuticals, where he was program leader for the drug delivery and generic Copaxone® programs and alliance manager for the Sandoz-Novartis collaboration. Prior to that, Dr. Hrkach was director of pulmonary formulations at Alkermes. Following his postdoctoral research with Professor Robert Langer at the Massachusetts Institute of Technology, he joined AIR, at its inception in 1998, and worked on large porous-particle technology-development programs. Dr. Hrkach is the author or coauthor of more than 25 scientific publications and 25 patents/applications in drug delivery and polymer chemistry. He received a B.S. in chemistry from the Philadelphia College of Pharmacy and Science and an M.S. and Ph.D. from Carnegie Mellon University in polymer science and chemistry, respectively.

**Ali Javey** is an assistant professor of electrical engineering and computer sciences at the University of California (UC)-Berkeley and a principal investigator in the Materials Sciences Division at Lawrence Berkeley National Laboratory. After receiving a Ph.D. in chemistry from Stanford University, Dr. Javey joined the faculty at UC-Berkeley in 2005. On leave for the 2005–2006 academic year, he was a junior fellow of the Harvard Society of Fellows. His research interests encom-

pass the fields of chemistry, materials science, and electrical engineering, with a focus on the integration of synthetic nanomaterials for technological applications, including high-performance nanoelectronics, flexible circuits and displays, and novel electronic sensors. He has published more than 30 papers in peer-reviewed journals, such as *Nature*, *Nature Materials*, *Physical Review Letters*, *Journal of the American Chemical Society*, and *Proceedings of the National Academy of Sciences of the United States of America*. His scientific awards include the Peter Verhofstadt Fellowship from the Semiconductor Research Corporation and the Graduate Student Gold Award from the Materials Research Society.

**John D. Lee** is a professor in the Department of Mechanical and Industrial Engineering at the University of Iowa and director of human-factors research at the National Advanced Driving Simulator. He is also affiliated with the Department of Neurology, Public Policy Center, Injury Prevention Research Center, and Center for Computer-Aided Design. His areas of research include the safety and acceptance of complex human-machine systems, the mediation of attention by technology; the building of trust in technology; advanced driver-assistance systems; and the causes and mitigation of driver distraction. Dr. Lee is coauthor of *An Introduction to Human Factors Engineering* (Prentice-Hall, 2003) and author or coauthor of more than 170 articles. He received the Ely Award for best paper in the journal *Human Factors* (2002) and the Best Paper Award in *Ergonomics* (2005); he is also a Donald E. Bently Faculty Fellow. Dr. Lee is a member of the National Research Council Committee on Human Factors and has served on several committees of the National Academy of Sciences. He is a member of the editorial boards of *Cognitive Engineering and Decision Making; Cognition, Technology and Work*; and *International Journal of Human Factors Modeling and Simulation* and associate editor of *Human Factors* and *IEEE-Systems, Man, and Cybernetics.*

**Joseph C. Martz** has been at Los Alamos National Laboratory (LANL) since he first arrived as a summer intern in 1983. From 1986 to 1990, he conducted research for his dissertation on the plasma processing of plutonium. After receiving his Ph.D. from UC-Berkeley, he returned to LANL, where he headed several projects on plutonium storage, including one that led to the nationwide mandate to stabilize stored nuclear materials, known as 94-1. In 1994, one of the youngest group leaders in Los Alamos history, he took charge of pit operations at TA-55. In 1997, he was program manager for weapon materials and enhanced surveillance, a position he held for nearly seven years, during which he led a number of special projects, including the Octave test series. From 2003 to June 2005, Dr. Martz was deputy division leader for X-Division, the principle nuclear weapons-design division. Subsequently, he was asked to lead the Reliable Replacement Warhead Program as the project director. An internationally recognized expert in

weapons materials and plutonium aging, Dr. Martz remains active in research and is a frequent consultant and contributor to intelligence programs.

**Samir Mitragotri** has been a professor of chemical engineering and professor of biomolecular science and engineering at the University of California-Santa Barbara since 2000. His research is on novel drug-delivery methods based on an understanding of control transport processes. He has received numerous awards, including the American Institute of Chemical Engineers Allan P. Colburn Award (2005), Controlled Release Society Young Investigator Award (2008), Ebert Prize from the American Pharmaceutical Association (1996), and the TR100 Young Innovator Award (1999). He is on the editorial boards of *Experimental Medicine and Biology* and *Journal of Controlled Release*. Dr. Mitragotri received a B.S. in chemical engineering from the University of Bombay and a Ph.D. in chemical engineering from the Massachusetts Institute of Technology.

**Steven D. Nixon** was, until recently, director of science and technology for the Office of the Director of National Intelligence (ODNI DS&T), one of eight legislated positions in ODNI, which oversees the science and technology activities of the 16 agencies in the U.S. intelligence community. Previous to that, he was deputy director of science and technology and first acting director of the newly established Intelligence Advanced Research Projects Activity. For the previous ten years, he was a member of the professional staff of the House Defense Appropriations Subcommittee, where he was responsible for reviewing a wide variety of military and intelligence research, development, and acquisition programs. In 2008, the DNI awarded him the National Intelligence Career Achievement Medal, and in 2005 the *National Journal* named him to the "Hill 100." Also in 2005, he was designated by *Space News* one of the top ten individuals "making a difference" in space. Prior to working with Congress, Mr. Nixon was a senior civilian analyst for the Department of the Navy. He received B.S. degrees in electrical engineering and mathematics from the University of Kansas and an M.A. in national security studies from Georgetown University. He is currently a consultant on national security matters in Washington, D.C.

**Mihrimah Ozkan,** associate professor in the Department of Electrical Engineering, University of California-Riverside (UCR), conducts interdisciplinary research on nanotechnology. She received a Ph.D. from UC-San Diego and an M.S. from Stanford University and has more than four years of experience in industry at Applied Materials, Analog Devices, and IBM Almaden Research Center. Her multidisciplinary research has focused on the developments of post-CMOS fabrication and integration methods for future electronics, hybrid organic-inorganic solar platforms, and "smart" nanoparticles for cancer therapeutics. Dr. Ozkan is the recipient of the Young Investigator Award from the U.S. Army (2006), Distinguished Engineering Educator of the Year Award from the National Engineers

Council (2006), Regents Faculty Excellence Award (2006, 2004, 2002), Emerging Scholar Award from the American Association of University Women (2005), invited participant to the National Academies Keck Future Initiatives Conference (2005), Visionary Science Award from the BioMEMS and Biomedical Nanotechnology Conference (2003), and Achievement in Technical Ingenuity Award from the Inland Empire Economic Partnership (2003). She is a member of the FCRP Center on Functional Engineered Nanoarchitectonics, National Science Foundation Center for Hierarchical Manufacturing, National Cancer Institute Nanotumor Center, U.S. Department of Defense Center of Nanoscience for Innovation in Defense, and the UCR Center for Nanoscale Science and Engineering. She is a member of the editorial boards of the *Journal of Sensors and Actuators B* and the *Journal of Biomedical Microdevices* and principal editor of *Micro and Nano Technologies for Genomics and Proteomics* (Springer, 2006). She has published approximately 150 journal papers, conference proceedings, and book chapters and holds more than 25 patent disclosures and 8 U.S. patents.

**Daniel W. Pack,** an associate professor of chemical and biomolecular engineering at the University of Illinois at Urbana-Champaign (UIUC), is affiliated with the Department of Bioengineering, the Center for Nanoscale Science and Technology, the Beckman Institute, and the Institute for Genomic Biology at UIUC. He completed a postdoctoral fellowship at the Massachusetts Institute of Technology and has received numerous awards, including the Multi-Year Faculty Achievement Award (2007), Xerox Award for Faculty Research (2008), and a National Science Foundation CAREER Award (2002). He received a B.S. from UIUC and a Ph.D. from the California Institute of Technology, both in chemical engineering.

**A.D. Romig, Jr.**, executive vice president and deputy laboratories director for Integrated Technologies and Systems and interim chief operating officer at Sandia National Laboratories, began his career there in 1979. His current responsibilities include leadership and management of development and engineering activities in support of military technology; proliferation prevention; technology assessment; counterintelligence; energy science, resources, conservation, and infrastructure assurance; and homeland security. Dr. Romig is a member of the National Academy of Engineering (NAE) and is active on a number of NAE/National Research Council committees and boards. He is a fellow of the American Association for the Advancement of Science and the Metals, Minerals and Materials Society (TMS), a Fellow and former president of ASM International (formerly American Society for Metals), and a Senior Member of IEEE. Dr. Romig is also on the boards of Atomic Weapons Establishment Management Limited and Technology Ventures Corporation, a member of the Council on Foreign Relations and the Intelligence Science Board, and has served on study committees of the Defense Science Board. Dr. Romig also serves on the board of HydroGen, LLC, and the board of directors of MIND Institute, a not-for-profit neuroscience company. He is the

recipient of the 2005 National Materials Advancement Award from the Federation of Materials Societies, the 2005 Acta Materialia Inc. J. Herbert Hollomon Award, and the 2003 ASM-TMS Distinguished Lecturer in Materials and Society. For his pioneering work in analytical electron microscopy and solid-state diffusion, Dr. Romig received the Burton Medal from the Microscopy Society of America (1988); the K.F.J. Heinrich Award from the Microbeam Analysis Society (1991); the ASM Silver Medal for Outstanding Materials Research (1992); and the Acta Metallurgica International Lectureship (1993–1994). He received his B.S., M.S., and Ph.D. in materials science and engineering from Lehigh University in 1975, 1977, and 1979, respectively.

**Jeffrey J. Welser,** on assignment from IBM Corporation, is director of the Nanoelectronics Research Initiative (NRI), a subsidiary of the Semiconductor Research Corporation. The goal of NRI, which supports university-based research on future nanoelectronics, is the development of logic devices capable of scaling beyond the limits of the CMOS transistor in the 2020 time frame. Dr. Welser joined IBM Research Division after receiving his Ph.D. in electrical engineering from Stanford University in 1995. At IBM, he worked on a variety of novel devices, including nanocrystal and quantum dot memories, vertical-FET DRAM, and silicon-based optical detectors; eventually he took over management of the Novel Silicon Device Group. At the time, he was also an adjunct professor at Columbia University. In 2000, he moved to IBM Technology group headquarters, and in 2001, he joined the Microelectronics Division as project manager for the high-performance CMOS device-design groups. In late 2003, he became director of Next-Generation Technology Components, and in 2006, he took on his current role at NRI. He is now based at the IBM Almaden Research Center in San Jose, California.

**Nikolai Zhitenev** is a project leader at the National Institute of Standards and Technology Center for Nanoscale Science and Technology. He received an M.Sc. in physics from the Moscow Institute of Physics and Technology and a Ph.D. in condensed-matter physics from the Russian Institute of Solid State Physics in 1991. He has worked at the Max Planck Institute for Solid State Physics, Massachusetts Institute of Technology, and at Bell Labs. His research focus is on electronic properties of novel materials in nanoscale devices.

# Program

**NATIONAL ACADEMY OF ENGINEERING**

2008 U.S. Frontiers of Engineering Symposium
September 18–20, 2008

Chair: Julia M. Phillips, Sandia National Laboratory

**DRUG DELIVERY SYSTEMS**

Organizers: William J. Grieco and Efrosini Kokkoli

*Recent Developments in Needle-Free Drug Delivery*
Samir Mitragotri

*Targeted Polymeric Nanotherapeutics*
Jeff Hrkach

*Polymer Technology for Gene Therapy*
Daniel W. Pack

*Traceable Drug Delivery: Lighting the Way with Qdots*
Xiaohu Gao

***

## EMERGING NANOELECTRONIC DEVICES

Organizers: Jia Chen and Victor Zhirnov

*The Quest for the Next Information-Processing Technology*
Jeffrey J. Welser

*Molecular and Polymer Nanodevices*
Nikolai Zhitenev

*Roll Printing of Crystalline Nanowires for Integrated Electronic and Sensor Arrays*
Ali Javey

*The Role of DNA in Nanoarchitectonics*
Mihrimah Ozkan

***

## COGNITIVE ENGINEERING

Organizers: Barrett S. Caldwell and Kim Vicente

*Cognitive Engineering: It's Not What You Think*
Stephanie Guerlain

*Driving Attention: Cognitive Engineering in Designing Attractions and Distractions*
John D. Lee

*Human Reliability Analysis in Cognitive Engineering and System Design*
Ronald Laurids Boring

*Cognitive Engineering Applications in Health Care*
Ann M. Bisantz

***

## UNDERSTANDING AND COUNTERING THE PROLIFERATION OF WEAPONS OF MASS DESTRUCTION

Organizers: J. Scott Goldstein and Gregory A. Hebner

*U.S. National Security in New Times*
Steven D. Nixon

*Combating Weapons of Mass Destruction: Translating Strategic Guidance into Actionable Solutions*
Charles L. Beames

*Nuclear Deterrence in the 21st Century: The Role of Science and Engineering*
Joseph C. Martz

***

## DINNER SPEECH

*Energy Policy and the Role of Technology in National Security*
A.D. Romig, Jr.

# Participants

**NATIONAL ACADEMY OF ENGINEERING**

2008 U.S. Frontiers of Engineering Symposium
September 18-20, 2008

Bryan Adams
Research Program Manager
Home Robots
iRobot Corporation

Holger Amort
Research Scientist
Arkema, Inc.

Jose Andrade
Assistant Professor
Department of Civil and Environmental Engineering
Northwestern University

Shelley Anna
Assistant Professor
Department of Mechanical Engineering
Carnegie Mellon University

Ralph Anthenien
Program Manager, Propulsion and Energetics
Mechanical Sciences Division
Army Research Office

Sasan Bakhtiari
Electrical Engineer/Section Manager
Nuclear Engineering Department
Argonne National Laboratory

Alexandre Bayen
Assistant Professor
Departments of Systems Engineering and Civil and Environmental Engineering
University of California, Berkeley

Charles Beames**
Chief, Intelligence, Surveillance, and Reconnaissance Division
Space and Intelligence Office
OUSD/AT&L
Department of Defense

Igor Benko
Researcher
Sun Microsystems Laboratories
Sun Microsystems

Ann M. Bisantz**
Associate Professor
Department of Industrial and Systems Engineering
University at Buffalo (SUNY)

Nicole Blute
Senior Project Engineer, Municipal
Malcolm Pirnie, Inc.

Alexandria Boehm
Assistant Professor
Department of Civil and Environmental Engineering
Stanford University

Joshua Boltz
Senior Process Engineer
CH2M HILL

Josh Bongard
Assistant Professor
Department of Computer Science
University of Vermont

Ronald L. Boring**
Senior Member of the Technical Staff
Risk and Reliability Analysis Department
Sandia National Laboratories

Donald Bott
Manager, Marking Elements & Integration Laboratory
Xerox Innovation Group
Xerox Corporation

Maggie Bynum
Research Scientist
Microfluidics and Biosensors Department
Agilent Laboratories

Barrett S. Caldwell*
Associate Professor
School of Industrial Engineering
Purdue University

Kirk Cameron
Associate Professor
Department of Computer Science
Virginia Tech

Jonathan Caulkins
Professor
Heinz School of Public Policy and Management
Carnegie Mellon University

Jia Chen*
Research Staff Member
IBM T.J. Watson Research Center

*Organizing Committee
**Speaker

Leo Chiang
Senior Research Specialist
Analytical Sciences, Core R&D
Dow Chemical

Mung Chiang
Associate Professor
Department of Electrical Engineering
Princeton University

Tushar Choudhary
Senior Scientist
ConocoPhillips Company

Catherine Corrigan
Corporate Vice President
Biomechanics Practice
Exponent, Inc.

Levent Degertekin
Professor
School of Mechanical Engineering
Georgia Institute of Technology

Tim Denison
Senior IC Technology and Development Manager
Medtronic, Inc.

Christopher Doerr
Distinguished Member of Technical Staff
Advanced Photonic Devices
Alcatel-Lucent/Bell Labs

Amy Duwel
MEMS Group Leader
Advanced Hardware
Charles Stark Draper Laboratory

Thomas Epps
Assistant Professor
Department of Chemical Engineering
University of Delaware

Malcolm Fabiyi
Group Leader, Research and Development
General Industries - Environmental Research
Praxair, Inc.

Roy Fielding
Chief Scientist
Research and Development
Day Software

Tim Frazier
Director
Cummins Technical Center
Cummins Inc.

Xiaohu Gao**
Assistant Professor
Department of Bioengineering
University of Washington

Chris Gearhart
Technical Leader
Fuel Cell/Hybrid Electric Vehicles
Ford Motor Company

Kathleen Gillespie
Integration Engineer, Energy Storage Systems
Hybrid Vehicle Integration
General Motors

J. Scott Goldstein*
Vice President of Technology
Chief Scientist of the Intelligence, Security, and Technology Group
SAIC

Alex Gorokhov
Principal Engineer/Manager
Corporate Research & Development
Qualcomm Inc.

Michael Graham
Harvey D. Spangler Professor
Department of Chemical and Biological Engineering
University of Wisconsin, Madison

James Gregory
Assistant Professor
Department of Aerospace Engineering
Ohio State University

William J. Grieco*
Vice President of Engineering Programs
PetroAlgae

Stephanie Guerlain**
Associate Professor
Department of Systems and Information Engineering
University of Virginia

Seth Guikema
Assistant Professor
Department of Geography and Environmental Engineering
Johns Hopkins University

Mitra Hartmann
Assistant Professor
Departments of Biomedical and Electrical Engineering
Northwestern University

Gregory Hebner*
Manager
Lasers, Optics, Remote Sensing, Plasma Physics, and Complex Systems Department
Sandia National Laboratories

Katherine Herrick
Program Manager
Systems Engineering/Advanced Technology
Raytheon

Jeff Hrkach**
Vice President
Pharmaceutival Sciences
BIND Biosciences

Inseok Hwang
Assistant Professor
Department of Aeronautics and Astronautics
Purdue University

Michael Isard
Senior Researcher
Microsoft Research

Paul Jacob
Associate
Offshore/Riser Systems
MMI Engineering, Inc.

Hamid Jafarkhani
Professor
Department of Electrical Engineering and Computer Science
University of California, Irvine

Ali Javey**
Assistant Professor, Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Principal Investigator, Lawrence Berkeley National Laboratory

Zhenhua Jiang
Assistant Professor
Department of Electrical and Computer Engineering
University of Miami

Vanja Josifovski
Principal Research Scientist
Yahoo! Research
Yahoo!

Annie Kammerer
Senior Seismologist/Seismic Engineer
Office of Nuclear Regulatory Research
Nuclear Regulatory Commission

Paul Kenis
Associate Professor
Departments of Chemical and Biomolecular Engineering
University of Illinois at Urbana-Champaign

Brian Kirby
Assistant Professor
Sibley School of Mechanical and Aerospace Engineering
Cornell University

Efrosini Kokkoli*
Assistant Professor
Department of Chemical Engineering and Materials Science
University of Minnesota

Carlo Kosik Williams
Senior Research Scientist
Science and Technology Division
Corning Incorporated

Mayuresh Kothare
Associate Professor
Department of Chemical Engineering
Lehigh University

Eva K. Lee
Associate Professor
School of Industrial and Systems Engineering
Director, Center for Operations Research in Medicine and Healthcare
Georgia Institute of Technology

John D. Lee**
Professor
Department of Mechanical and Industrial Engineering
University of Iowa

Robert Leland
Director
Computing and Network Services Center
Sandia National Laboratories

Guifang Li
FPCE Professor
CREOL, The College of Optics & Photonics
University of Central Florida

J. Christopher Love
Texaco-Hangeldorf Career Development Professor
Department of Chemical Engineering
Massachusetts Institute of Technology

Reece Lumsden
Project Manager
777 Electrical Design
Boeing Company

Zhengxiang Ma
Principal Engineer
Wireless Research
FutureWei Technologies, Inc.

Ellen Mangan
Senior Usability Team Lead
Signal Processing and Communications
MathWorks, Inc.

Joseph Martz**
Senior Staff Member and Advisor
Principle Associate Directorate for Nuclear Weapons
Los Alamos National Laboratory

Russell Meller
Hefley Professor of Logistics and Entrepreneurship
Department of Industrial Engineering
University of Arkansas

Karen Miu Miller
Associate Professor
Department of Electrical and Computer Engineering
Drexel University

Samir Mitragotri**
Professor
Departments of Chemical Engineering and of Biomolecular Science and Engineering
University of California, Santa Barbara

Raymond Moran
Senior Project Manager
Construction Services
PB Americas, Inc.

Vivek Narasayya
Senior Researcher
Data Management and Exploration Group
Microsoft Research

Steven D. Nixon**
Independent Consultant on National Security Issues

Patrick O'Mara
Senior Transportation Engineer
Transportation & Infrastructure - Planning
STV Group, Inc.

Mihri Ozkan**
Associate Professor
Department of Electrical Engineering
University of California, Riverside

Daniel W. Pack**
Associate Professor
Department of Chemical and Biomolecular Engineering
University of Illinois at Urbana-Champaign

Matteo Pasquali
Associate Professor
Department of Chemical Engineering
Rice University

Julia M. Phillips*
Director
Physical, Chemical, and Nano Sciences Center
Sandia National Laboratories

Michael Picciolo
Senior Research Scientist
ASSET Business Unit
SAIC

Mark R. Prausnitz
Professor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Amy Pruden
Associate Professor
Department of Civil and Environmental Engineering
Virginia Tech

Naren Ramakrishnan
Associate Professor and Faculty Fellow
Department of Computer Science
Virginia Tech

Christopher Schuh
Associate Professor
Department of Materials Science and Engineering
Massachusetts Institute of Technology

Jeffrey Scruggs
Assistant Professor
Civil and Environmental Engineering
Duke University

Jayshree Seth
Division Scientist
Industrial Adhesives and Tapes Division
3M

Marcus Shaw
Senior Project Leader
Architecture and Design Subdivision
Aerospace Corporation

Andrew Singer
Director, Technology Entrepreneur Center
Professor, Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign

James Singletary
Senior Research Engineer
Advanced Fibers Systems
DuPont

Eric Smith
Nuclear Engineer
National Security Directorate
Pacific Northwest National Laboratory

Narayan Srinivasa
Senior Research Scientist
Information and System Sciences Office
HRL Laboratories LLC

Shaun Sullivan
Systems and Performance Engineer
Applications Engineering
Ingersoll Rand

Mark T. Swihart
Associate Professor
Department of Chemical and Biological Engineering
University at Buffalo (SUNY)

Jose Tabora
Senior Principal Scientist
Bristol Myers Squibb

Ichiro Takeuchi
Associate Professor
Department of Materials Science and Engineering
University of Maryland

Hong Tang
Assistant Professor
Departments of Electrical and Mechanical Engineering
Yale University

Nelson Tansu
Peter C. Rossin Assistant Professor
Department of Electrical and Computer Engineering
Center for Optical Technologies
Lehigh University

Jason Trachewsky
Senior Technical Director and Broadcom Fellow
Broadcom Corporation

Andrew Tsourkas
Assistant Professor
Department of Bioengineering
University of Pennsylvania

Stephen Uftring
Assistant Group Leader
Lincoln Laboratory
Massachusetts Institute of Technology

Omar Varela
Research Specialist
ExxonMobil Upstream Research Company

Andrew M. Weiner
Scifres Family Distinguished Professor of Electrical and Computer Engineering
Purdue University

Jeffrey Welser**
Director
IBM and SRC Nanoelectronics Research Initiative
IBM Almaden Research Center

Peter Winzer
Distinguished Member of Technical Staff
Alcatel-Lucent/Bell Labs

Jihui Yang
Staff Research Scientist
Materials and Processes Laboratory
General Motors Research & Development Center

Ji-Cheng Zhao
Associate Professor
Department of Materials Science and Engineering
Ohio State University

Victor Zhirnov*
Research Scientist
Semiconductor Research Corporation

Nikolai Zhitenev**
Project Leader
Center for Nanoscale Science and Technology
National Institute of Standards and Technology

Michelle Zhou
Research Scientist and Manager
IBM T.J. Watson Research Center

**Dinner Speaker**

Alton D. Romig, Jr.
Executive Vice President, Deputy Laboratories Director for Integrated Technologies and Systems, and Interim Chief Operating Officer
Sandia National Laboratories

**Guests**

Adnan Akay
Division Director
ENG/CMMI
National Science Foundation

Cesar Castro
Director, Business Development
Institute for the Future

Joe Cecchi
Dean, School of Engineering
University of New Mexico

William Hayden
Vice President
The Grainger Foundation

Sohi Rastegar
Director
Emerging Frontiers in Research and Innovation
National Science Foundation

William Rees
Deputy Under Secretary of Defense (Laboratories and Basic Sciences)
Department of Defense

Phillip Szuromi
Supervisory Senior Editor
Science Magazine

**National Academy of Engineering**

Charles M. Vest
President

Lance A. Davis
Executive Officer

Janet Hunziker
Senior Program Officer

Gin Bacon
Program Associate