Preponderance of Evidence
Judging What to Do About Ephedra

By Paul G. Shekelle, Margaret Maglione, and Sally C. Morton
Get the Big Picture

RAND Review covers the big issues with an eye for the important details.
Message from the Editor

Rarely does research reach into your medicine cabinet and prompt you to take another look at those pills that you might have thought were harmless just because they were “natural.” Our cover story may do just that.

Paul Shekelle, Margaret Maglione, Sally Morton, and others have illuminated the debate about the safety and efficacy of the herb ephedra, which millions of Americans have consumed with the intent of losing weight, building muscle, or boosting energy. The research comes in the wake of several highly publicized deaths of people who had apparently ingested the herb shortly beforehand.

These deaths, combined with the research, have prodded some lawmakers to reconsider the wisdom of a 1994 piece of legislation that eased the regulatory standards on herbs and other dietary supplements. The legislation also made it tougher for researchers to gather definitive data on ephedra. Our cover story explains how the research team built a compelling case from the medical evidence available on the herb despite the absence of conclusive clinical data.

Our next feature story outlines the contrasting lessons learned from two very different educational interventions. Susan Bodilly, Brian Gill, and others describe the results of a business-driven intervention that had a poor community plan and of a community-driven intervention that had a poor business plan. Both interventions met with mixed success. The lessons learned in each case could help future efforts meet with greater success.

To supplement direct military and police actions against terrorism, Paul Davis and Brian Jenkins describe principles for a sustainable “systems approach” that recognizes the many different types of actors in terrorist organizations and that seeks to influence them in a broad range of ways. The authors also discuss some of the most vexing challenges involved in implementing the approach.

From RAND Europe, Caroline Wagner warns that global anxiety about weapons of mass destruction should not divert attention away from the proliferation of “weapons of mass creation.” These weapons, associated with genetic manipulation, could also threaten society. Her message reaches beyond the medicine cabinet and into the womb of every expectant mother.

—John Godges

4 News
- Survival tips from the wild kingdom
- $50 million needed for landmine detection
- Why it’s smart to vaccinate health care workers against smallpox
- Unexpected benefits of drug prevention
- Economics of age structures
- Gaping hole in dental care
- Smaller gifts for higher education
- Ripple effects of rising health insurance premiums

8 Perspectives
For Better, Not Worse
Low-income women may no longer be tolerating abuse

10 Hazardous Haste
Warning systems themselves could be dangerous

12 The Influence Component of Counterterrorism
A systems approach
By Paul K. Davis and Brian Michael Jenkins

16 Cover Story
Preponderance of Evidence
Judging what to do about ephedra
By Paul G. Shekelle, Margaret Maglione, and Sally C. Morton

22 Crash Courses
Hard lessons from educational interventions
By Susan J. Bodilly, Brian P. Gill, Mark Berends, Sheila Nataraj Kirby, Jacob W. Dembosky, and Jonathan P. Caulkins

30 Commentary
Beware the weapons of mass creation
By Caroline S. Wagner

On the Cover
Wild ephedra grows in Oljeto Canyon, Utah.
News

Animal Behavior Offers Clues for Military Deception

The animal kingdom is becoming a wellspring of insight for researchers hoping to develop novel military technologies and techniques.

In a recent RAND report, researchers examine deceptive techniques employed by various animals and suggest leveraging these techniques for use on the modern human battlefield. The researchers call for increased cultivation of deception techniques to support the U.S. military, notably for urban operations.

The scientific literature on animal and plant deception—where ruses of virtually infinite variety are applied to offense, defense, and intelligence gathering—is often richer and more rigorous than corresponding literature on military deception, the report notes.

All organisms seek accuracy in their perceptions, because survival depends on it. The corollary holds true: Causing inaccuracy in the perceptions of an adversary is advantageous. This is the essence of deception. One example is the “false head” or “eyespot” used by many insects to lure, frighten, or confuse enemies. An effective technique in the animal kingdom, the false head already has a parallel in human warfare: the false cockpit painted on the underside of the Canadian CF-18 Hornet air-to-air fighter.

This inexpensive bit of deception can produce a momentary—but decisive—advantage for the pilot in combat.

“Deception techniques are pervasive and enduring throughout the natural world, with a single species often employing more than one type of ruse for greater effect,” said Scott Gerwehr, primary author of the report.

“Given how intolerant natural selection is of wasted effort, this suggests that deception is an adaptation as valuable as armor, speed, or firepower—and should be cultivated in our soldiery on both the technology and training sides.”

For more information: Unweaving the Web: Deception and Adaptation in Future Urban Operations (RAND/MR-1495-A).

Bigger Effort Needed to Remove Landmines

A major research and development program costing about $50 million over five to eight years is needed to sharply accelerate efforts to remove landmines that kill thousands of civilians each year in 90 nations, according to a RAND report.

“There is a desperate need for better landmine detection equipment,” said Jacqueline MacDonald, an engineer and coauthor of the report. “Technology is available to create better tools to remove landmines, but nothing will be developed unless there is investment in a well-organized, focused research program.”

The report said research is needed to develop new technology that can replace the World War II-era equipment—the mainstay of worldwide efforts to remove landmines. Researchers cited the need for a new generation of landmine detectors that would be more accurate and reliable to speed landmine removal.

Today’s landmine detection equipment is primitive, relying on technology that results in a high number of false alarms, according to researchers. Landmine detectors used today operate via a technology that is unable to distinguish landmines from other metallic materials—by far the greatest limitation of the process.

Current levels of funding are insufficient, says the report. In 2002, the United States invested only $2.7 million for research and development on handheld detectors for humanitarian de-mining. The report concludes that no single mine detection technology exists that operates effectively against all mine types in all settings.

For more information: Alternatives for Landmine Detection (RAND/MR-1608-OSTP).
A “Prudent” Approach to Smallpox Vaccinations

Widespread vaccination of Americans against smallpox is too dangerous to justify unless the likelihood of a major biological attack on the United States is substantial, but it is prudent to vaccinate health care workers now, according to a RAND study published in The New England Journal of Medicine.

The researchers believe that the risk of vaccinating health care workers is justified, because they are particularly vulnerable during smallpox outbreaks. Health care workers could come into close contact with those sick with smallpox, often before the disease is recognized.

Moreover, “if health care workers themselves became sick, it could threaten our ability to keep our hospitals open and our health care system running,” said Samuel Bozzette, the study’s lead author and infectious disease specialist affiliated with RAND, the Veterans Affairs San Diego Healthcare System, and the University of California at San Diego. “Vaccinating health workers presents a modest risk and could pay many benefits.”

For this study, health care workers were defined to include everyone working in health and related services who would come into direct contact with sick people. This includes doctors, nurses, and others working in and around health care facilities, including paramedics, ambulance drivers, and hospital security guards.

Researchers developed a number of feasible smallpox attack scenarios, reviewed historical smallpox outbreaks, and developed a computer model that estimates the number of people who might be killed in each of the attacks under a number of policy options.

The researchers found that if nearly all 10 million health care workers in the United States were vaccinated against smallpox, an estimated 25 people would die. The table below indicates the number of health care workers and others who could die under various smallpox attack scenarios.

### Smallpox Scenarios: Predicting the Damage

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Deaths from smallpox vaccination (health care workers)</th>
<th>Deaths from smallpox (general public)</th>
<th>Total deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>If suicide attackers ride mass transit spreading the virus</td>
<td>If health care workers are vaccinated before an attack</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>If health care workers are vaccinated after an attack</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>If virus is sprayed in ventilation system of office building</td>
<td>If health care workers are vaccinated before an attack</td>
<td>25</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>If health care workers are vaccinated after an attack</td>
<td>1</td>
<td>298</td>
</tr>
<tr>
<td>If a low-impact* airport attack is successful in spreading the virus in airport terminals</td>
<td>If health care workers are vaccinated before an attack</td>
<td>26</td>
<td>2,192</td>
</tr>
<tr>
<td></td>
<td>If health care workers are vaccinated after an attack</td>
<td>26</td>
<td>2,731</td>
</tr>
<tr>
<td>If a high-impact* airport attack is very successful in spreading the virus in multiple airport terminals</td>
<td>If health care workers are vaccinated before an attack</td>
<td>49</td>
<td>43,852</td>
</tr>
<tr>
<td></td>
<td>If health care workers are vaccinated after an attack</td>
<td>55</td>
<td>54,643</td>
</tr>
</tbody>
</table>


*In the low-impact case, 5,000 people are infected. In the high-impact case, 100,000 people are infected.
News

School Drug Programs Benefit Society in Unexpected Ways

Two-thirds of the drug-related benefits from school-based drug prevention programs come from reduced tobacco and alcohol use, and the programs offer clear economic gains to society, according to a RAND study.

Only a third of the benefits involved the reduced use of illicit drugs, primarily cocaine. Reduced marijuana use accounted for a very small fraction of the benefits.

The benefits included lower health care costs, higher productivity, and lower costs to the criminal justice system.

“While we commonly think of these antidrug programs as targeted at street drugs, it turns out that most of the benefits that society receives are due to the cuts in the use of legal substances,” said Jonathan Caulkins, lead author of the report.

“It may make sense to view school-based drug education programs as health promotion efforts generally—or as investments in the human capital of the next generation of Americans—rather than as purely anticrime initiatives.”

Whereas the social cost of the typical school-based prevention program is on the order of $150 per student, researchers estimate that society receives $840 in benefits for each student who takes part. Even when using very conservative assumptions about the effectiveness of drug prevention programs, researchers estimate that the programs provide at least $300 in benefits for each participating student.


Reducing High Fertility Rates Can Yield Economic Benefits

Reducing high fertility rates can help nations create a population age structure that is more likely to produce economic benefits—but only if policies are put in place to ensure quality health care, education, and job opportunities, according to a RAND report.

“Population change exerts a powerful influence on economic well-being,” said David Bloom, co-author of the report. For decades, experts have debated the relationship between population dynamics and economic performance. The report concludes that the debate has overlooked the effects of age structure—the way the population is spread across age groups.

When fertility rates fall, so does the number of young dependents. Soon, the proportion of working-age people increases (see figure). With a larger share of working-age people, nations may see a substantial boost in income growth and savings. This phenomenon is known as the “demographic dividend.”

However, the demographic dividend is not automatic. For workers to be productive, they require education, jobs, and health care. Countries that fail to provide the necessary environment may face grim futures, Bloom said. These countries may face trouble as unemployment rises, the social fabric frays, and rising numbers of old people overwhelm the available resources.

Half of U.S. Kids Don't Get Proper Dental Care

Nearly 50 percent of U.S. kids don’t benefit from the recommended number of dental visits, and nearly 25 percent don’t benefit from the recommended number of preventive pediatric visits. These are the findings of a recent RAND study published in Pediatrics.

The American Academy of Pediatric Dentistry and the American Academy of Pediatrics recommend schedules for dental visits and preventive pediatric visits—schedules that vary depending on the age of the child. The study of nearly 36,000 children under 18 years of age revealed substantial gaps between the recommended levels of care and the care actually received.

Children who are uninsured, who have a parent who is less than college educated, or who are in poor health are most at risk for not receiving the recommended care. Publicly insured children are more likely to make the recommended number of pediatric visits.

The RAND study calls for improving access to dental care for children, especially among adolescents and children in poor general health. “Dental health is critical for raising healthy children, but many children aren’t getting adequate dental care,” said Mark Schuster, an author of the study.

“Policymakers need to address this gap when considering policies to promote child and adolescent health,” he said.

Contributions to Higher Education Decline

Contributions to colleges and universities in the United States dropped slightly during 2002—the first decline since 1988, according to the Council for Aid to Education, a subsidiary of RAND.

Private gifts to higher education declined 1.2 percent to $23.9 billion for the fiscal year that ended June 30, 2002. The drop was fueled primarily by a sharp decline in gifts from alumni, according to the survey, which has tracked giving to higher education for more than 50 years.

“The declining stock market and the weak economy are the two primary factors that contributed to the decline in contributions,” said researcher Ann Kaplan. “Unfortunately, this decline in giving comes at a time when higher education institutions are seeing an overall retrenchment from other income sources as well.”

Employees Take Hits When Health Insurance Premiums Rise

Employees respond to an increase in the price of their health insurance by reducing their take-home pay and other benefits, such as life insurance, disability and dental insurance, and retirement benefits, according to a RAND study published by the National Bureau of Economic Research.

“Our results suggest that rising health insurance prices reduce not only employees’ take-home pay but also their insurance purchases against a variety of other risks,” said Dana Goldman, the lead author of the study. “If health insurance and other insurance products continue to rise—and individuals continue to reduce their purchases of health insurance and other insurance products—that might leave them vulnerable to health, mortality, disability, and other significant risks in the long run.”

The study examined employee choices within a single large firm that offered a flexible benefits plan. Under such an arrangement, employees explicitly choose how to allocate compensation between cash and various benefits. Such plans cover about half of all workers in medium and large firms, and that proportion is growing.

The researchers found that a $1 increase in the price of health insurance leads employees with flexible benefits plans to increase their expenditures on health insurance by just 52 cents. The employees then pay about two-thirds of the increased expenditure out of their own wages and the remaining third by reducing other benefits, such as retirement.
SHOULD TAXPAYERS SUBSIDIZE MARRIAGE for the low-income? It’s a question that Andrew Cherlin of Johns Hopkins University thinks about a lot these days. He’s the coinvestigator for a multimillion-dollar study that attempts to define how physical and sexual abuse early in life affect the rate of marriage later in life. His study strives to shed light on various questions related to welfare reform.

The 1996 welfare reform law, known as the Personal Responsibility and Work Opportunity Reconciliation Act, “emphasizes marriage a lot,” said Cherlin. “In fact, it states that marriage is a fundamental good, that non-marital childbearing is a fundamental problem.”

But Cherlin has raised doubts about the supposed connection between a long-term marriage and a “healthy” marriage. He recently spoke to RAND about some preliminary findings of a new study he is conducting with Linda Burton, Tera Hurt, and Diane Purvin.

The Role of Abuse

Cherlin began his research by asking himself why there weren’t more long-term unions among economically disadvantaged adults. Why do the low-income have such a low rate of marriage?

Liberals have blamed the problem on a lack of jobs or a lack of decent-paying jobs. Conservatives have placed the blame on a decline in morality or character.

To assess the validity of these arguments, Cherlin’s study team knocked on 40,000 doors to find a representative mix of 2,000 families in the welfare-eligible population. The study focused on low-income neighborhoods in Boston, Chicago, and San Antonio. Most people in the survey were African Americans or Hispanics in their 20s or 30s. The predominant marital status was “not married” or “not cohabitating.”

There may be some truth to both the liberal and conservative arguments, but neither of them can explain very much about the problem of nonmarital childbearing, said Cherlin. Instead, he suggested that there could be an additional and perhaps increasingly important factor in the lives of low-income women: the acute affect of the traumatic experiences that these women start to have during childhood and continue to have into adulthood.

Empirical evidence from the study suggests that domestic abuse may be no more or less widespread than in the past, but domestic abuse is more of a determining factor in marital breakups now than it was 20, 30, or 50 years ago.

He cited two kinds of traumatic experiences in particular: sexual abuse and physical abuse. The sexual abuse could start in childhood. The physical abuse could start either in childhood or in adulthood. These experiences are often overlooked as potentially important reasons for the decline in marriage. Sexual abuse and physical violence are not new, said Cherlin, but they are now occurring in a changing social context.

Compared with a few decades ago, there’s greater acceptance of and feasibility in living outside of marriage today and having a child. There are also changing expectations for marriage, even among the poor, with different values about what a marriage should be.

The economic feasibility and the increased moral tolerance of living as a single adult with a child are giving women more discretion in choosing not to marry. Simultaneously, the women may be viewing relationships differently than they otherwise might have a few decades ago.

“What I’m not claiming is that there is more sexual abuse and physical violence than there was a few decades ago,” Cherlin emphasized. “There’s no evidence whatsoever that there is more than there was. I am claiming—in a speculative argument—that [the violence] is operating in a different climate where women can make different decisions and have more options than they used to.”

According to Cherlin, a high level of sexual abuse and physical violence beginning in childhood could lead women to a wariness toward men and to a desire to maintain control over the relationships.
“Based on the literature, I would suggest that for women who have experienced childhood sexual abuse, the consequence is not necessarily less time spent in [domestic] union. Rather, it becomes more brief unions, short-term transient unions that are informal and nonmarital in character.”

The literature on childhood sexual abuse “shows paradoxically that women who have been abused sexually as children, when they become adults, have more sexual partners rather than fewer sexual partners.”

Research on HIV and risky sex also shows that women who have been abused in childhood engage in riskier sexual behavior compared with those who have not been abused. The research further shows that women who have been abused report that sex is less pleasurable and more anxiety producing.

There appears to be a dynamic set up by childhood sexual abuse. It can result in more sex but more anxiety about it. “That, too, results in an adult pattern of brief or informal unions,” said Cherlin.

For people who are physically abused, especially among adults, there is some evidence of the pattern of transient or more numerous unions. But just as important is a general avoidance of men altogether.

Cherlin referred to a recent study of Mexican immigrant women from different generations. The first generation said that the most important thing to them in a relationship was respect. Even if a husband was unfaithful, that by itself was not grounds for divorce—as long as he treated his wife with respect.

The second and third generation spoke in different terms, alluding to something that they translated as “intimacy.” In an intimate relationship, it wasn’t enough for a man just to have respect for his wife. These women would view his infidelity as a very bad sign of something that could lead to the end of the relationship.

“In other words, the second and third generation were thinking of marriage in terms of companionship and intimacy and the kinds of things that the middle class takes for granted,” said Cherlin. “This seems to be a shift in the views of Mexican immigrants, and my guess is that it’s a shift for many other poor women in other groups as well.

“I think there is a very real shift from several decades ago in the U.S. and in much of the rest of the developed world in the way that people look at relationships. We have different standards now, and there are also different costs involved when those standards aren’t met.”

Mary Jones, left, of the Arkansas Coalition Against Domestic Violence, rallies support for funding domestic violence shelters. Jones, whose daughter was killed in a domestic violence incident, gathered with supporters on March 13 on the steps of the state capitol in Little Rock.
Hazardous Haste
Warning Systems Themselves Could Be Dangerous

IT’S A NIGHTMARE SCENARIO: Reports begin to surface that a handful of city residents have taken ill with an undetermined disease. Within days, thousands more people grow sick. People begin to die, especially the elderly and those with weakened immune systems. Soon, the source of the pathogen is identified: the city’s water source. Corrective measures are taken, and warning letters are sent to residents, but it’s too late. More than 400,000 residents are infected, 4,000 are hospitalized, and 100 die.

This scenario is not from the pages of an emergency response plan. It’s a true story, one that happened in Milwaukee only ten years ago. Such experiences are now being mined for lessons by policymakers, who, faced with growing fears of terrorism in the United States, are examining the role of existing early-warning systems to safeguard public health.

Some of the lessons now being learned are related to “risk communications.” Baruch Fischhoff, an expert on risk communications at Carnegie Mellon University, spoke recently at RAND about the inadequacy of current warning systems. He argued that effective warning systems involve four components: a technical analysis of the potential risks of issuing warnings, an establishment of priorities, a clear understanding of public beliefs and values, and an evaluation of the outcomes of warnings that are issued.

He discussed a study for the American Water Works Association Research Foundation (AWWARF). The AWWARF wanted to develop more-effective communications, informing people about the risks of contaminants in the water supply and the need to boil the water before drinking it.

One question posed by the project was whether improved boil-water notices could reduce mortality associated with ingestion of a tiny parasite called cryptosporidium—the same pathogen that caused the Milwaukee outbreak. Proper boiling of water kills the parasite, which is resistant to most chemical disinfectants. Cryptosporidium can infect the human intestines, leading to severe gastrointestinal problems, and can cause death among the elderly and immunocompromised populations.

Fischhoff’s research team created a model to predict the effects of various boil-water notices in the event that the parasite found its way into a city water supply. The analysis considered what would happen with the “perfect” boil-water notice, which reached all consumers and led to effective responses, compared with what would happen with the conventional boil-water notices. The team built the model on assumptions gleaned from interviews and previous research. To their surprise, there was absolutely no change in mortality. “We ran the model, and our perfect notice had no impact whatsoever,” said Fischhoff.

The research group eventually realized that notices were bound to be futile, because it takes up to a week to examine water cultures for cryptosporidium before any notices can be issued. “By the time you find the cryptosporidium in the water supply, the people who are vulnerable have already taken the hit,” said Fischhoff.

“We had a system that was relying on notices that were pointless.” Knowing this should cause us to “look for alternatives, such as investing in better cryptosporidium screening technology or making it feasible for susceptible populations to have continuous access to purified water.”

The same model would show, however, that warning the public could be effective with contaminants like the bacterium E. coli, which can be cultured more quickly. “If you’re worried about E. coli, it’s possible to create an effective boil-water notice that could significantly affect risk, because the notice could be sent out in time to make a difference.”

Duct Tape: Use for Taping Ducts
Shortly after the U.S. Department of Homeland Security raised the national terrorist attack threat level from yellow to orange on Feb. 7, a news report quoted a spokesperson from the Federal Emergency Management Agency. The spokesperson suggested that Ameri-
cans keep on hand a supply of plastic sheeting and duct tape to seal windows from biological or chemical agents. While such advice might seem logical, in reality it may offer no benefit, according to Fischhoff. He cited the speculation of Israeli colleagues, who wondered whether their use of duct tape during the 1991 Persian Gulf War had been pointless. At the time, they thought that chemical-laden missiles would have had to hit extremely close to a home to disperse lethal concentrations of a chemical weapon. Yet at such a close range, the explosion from the missile could have dislodged the duct tape seal anyway.

The Israeli government generally has a good reputation for communicating the truth, however harsh it is, said Fischhoff. Nonetheless, credibility requires a perception of both competence and honesty. In the case of the duct tape, the government may have failed to conduct an appropriate technical analysis of the efficacy of the tape. As a result, the authorities weakened their credibility. “Once you lose it,” he warned, “it’s really hard to go back.”

Americans, too, have a strong desire for “honest, accurate information—even if [it] worries people,” said Fischhoff. He cited national surveys that he and his colleagues conducted in late 2001 and 2002. Yet he fears that “here in the United States, we now have a risk communication system that lacks the commitment to creating and providing honest, accurate information in a comprehensive form. Without the technical analysis of risks and empirically evaluated communications, I think we are digging ourselves into a hole.”

For example, the difference between “dirty bombs” and “nuclear bombs” is apparently not clear to the public, even though the difference is huge and potentially easy to explain. “Although there are physicists deeply concerned about public welfare, they cannot [aid the public] alone, without properly developed communication.”

**Danger! Danger! Experts!**

It takes a cooperative and collaborative relationship between the experts and the public to create a warning system that works, said Fischhoff. One common pitfall for experts is focusing on problems that may not be the ones really facing people.

He emphasized that experts should seek to understand what the audience already knows and then build on that base of knowledge. At the same time, he warned against underestimating the public. “Don’t shortchange people’s ability to run their own lives by giving them lousy information or irrelevant information. People will often do sensible things if they get relevant information in a credible, comprehensible form.”

He sent a warning specifically to the audience convened at RAND. “As experts, we need to realize that we could be the enemy, that we might not be trusted. As a result, we bear a special responsibility to serve the public’s needs, so as to warrant credibility.”

He said there are several fields that have been wrestling with these issues for a long time. In both the environmental profession and the chemical industry, for example, a credible process of effective risk communication often takes place. But he said that the medical field lags behind, and homeland security is even further behind.

Periodic evaluations of warning systems are critical, he concluded. “If the system isn’t working—like the boil-water notices—then come up with a different system. Don’t just leave people with impossible problems and then hold them responsible for ineffective action.”

---

*Store clerk Greg Dunlap reaches for an Israeli-made M-15 gas mask at The Barracks military supply store in downtown Boston on Feb. 14. Store clerks sold ten of the masks in the previous week since the raising of the nation’s terror threat level from yellow to orange.*

---

*AP/WIDE WORLD PHOTOS/ELISE AMENDOLA*
Early in 2001, we were asked by the director of the Defense Advanced Research Projects Agency to work together with the Institute for Defense Analyses (IDA) in constructing a framework for thinking about the deterrence of terrorism, and al Qaeda in particular. At the time, we and other Americans had in mind rather different and unpleasant outcomes for Osama bin Laden and associates (death or capture). However, the request was a good one because the United States needs a counterterrorism strategy to supplement direct action and direct defense. Consequently, the RAND and IDA teams worked to sketch an appropriate framework, one that would be at once compact and meaningful. The result was to emphasize four principles as follows.

**Four Guiding Principles**

Our primary strategic principle is that the United States should move beyond the concept of deterrence and toward influence as the appropriate complement to direct military action. Deterrence—or causing an adversary to desist from an action by threatening an even stronger counteraction—is simply too limiting. It may be impossible, for example, to deter members of al Qaeda by threatening to imprison or kill them. However, there are many ways to influence an adversary, only one of which is deterrence. Other ways include co-optation and inducement on the nicer-than-deterrence end, and crushing defeat on the harsher-than-deterrence end (such defeats should, of course, help deter future events). This spectrum of influences was motivated in part by our reading of history, which reminded us that even murderous terrorists have been of different types, with some of them being incorrigible and others moving on to become part of civil society when conditions changed. If there is a model for thinking about such matters, it is that terrorism should be sharply defeated, but that conditions must then change or it may very well arise again.

This concept of a spectrum of influences and an eye on both near-term and longer-term effects lead to our second guiding principle: Terrorist groups are not simply single entities; rather, they are systems, with diverse elements, many of which may be amenable to some influences but not others. Even though zealous leaders may not be deterrable, others who support a terrorist organization with money, logistics, and sanctuary may very well be. Indeed, many such people have a good deal to lose. The larger terrorist system, then, includes merchants, suppliers, heads of state, population segments, and religious leaders sympathetic to the terrorist network. Only some parts of the system can be
deterred in the classic sense. However, each part of the system can be influenced in myriad ways.

Our third principle is that because al Qaeda has no well-defined “center of gravity,” the United States should conduct a broad-front, sustainable campaign against the many components of the terrorist system. With no way of knowing in advance which part of the effort may prove to be the most successful, the United States needs to wage a simultaneous campaign along many fronts. This is quite different from ordinary warfare and may seem inefficient, but there is no choice.

Our fourth principle is that to sustain this type of campaign over many years, the campaign must be persuasive, morally high-minded, and consistent with enduring American values. (Americans, and other people in fear of survival, may stretch or violate their values temporarily, but the enduring values reemerge.) The war on terrorism should be characterized not only by manifest strength, purpose, and determination, but also by a moral validity readily apparent to U.S. citizens and to our counterterrorism partners worldwide. Moreover, as signaled above, the campaign should balance short-term efforts to destroy terrorists with longer-term efforts to diminish the public appeal and power of terrorism and to address underlying problems.

Three Troublesome Challenges

Developing and executing this strategy will not be easy. To succeed, the United States will need to overcome many vexing problems. We highlight three: (1) deterring the use of weapons of mass destruction, (2) persuading regional allies to act, and (3) maintaining American values at home and abroad.

**Deterring the Use of Weapons of Mass Destruction.** America’s most worrisome national security concern is the specter of catastrophic terrorism involving weapons of mass destruction (WMD). We suggest two approaches beyond those already being taken. The first is to announce credibly that any state or nonstate organization that even tolerates the acquisition of WMD by terrorists within its borders will be subject to U.S. military action.

The second approach is quite different and controversial. Deterrence of the use of biological weapons could be greatly enhanced if everyone in the Middle East discussed the ramifications of biological weapons sufficiently so as to conclude—from their own common sense—that biological warfare—once begun—could not be contained. Therefore, the United States should publicize the fact that a major bioterrorism attack involving a highly contagious disease such as smallpox would, because of international travel and inevitable retaliation by one party or another, almost certainly come back to haunt the Middle East and might even result in a global pandemic. Would-be terrorists and their supporters must be convinced that such weapons cannot be “controlled” or used with impunity.

Despite major difficulties, the U.S. public health system would be able to cope with an outbreak of disease associated with a bioterrorist attack. European public health systems would cope as well. But with frail public health institutions and limited medical capabilities, the world’s poorer nations, particularly those in the Middle East, would suffer enormously, perhaps losing substantial portions of their populations.

In short, U.S. information strategy should aim to induce the populations who support terrorists to recognize the likely self-defeating consequences of biological warfare. Potential advocates or tolerators of such warfare should conclude that this is “not a game we should even think about playing.”

**Persuading Regional Allies to Act.** Our systemic approach to counterterrorism distinguishes among the various kinds of adversaries who support terrorism. We decompose the terrorist system into several different types of actors, what they hold dear, and what could threaten or influence them (see table). Identifying these actors and the levers of influence, however, is the easy part. The hard part is making the proposed actions actually happen, especially when many of the actions need to be taken by states from which terrorists come or in which they reside. Clearly, many of the actions proposed in our table need to be performed by other countries.

America’s allies in Europe began vigorous crackdowns on terrorists immediately after Sept. 11, 2001, and have reportedly cooperated closely with U.S. authorities since. Middle Eastern governments—notably Saudi Arabia, Egypt, and Pakistan—must also
<table>
<thead>
<tr>
<th>What the Participants in Terrorism Hold Dear</th>
<th>What the United States and Its Allies Might Do</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Terrorist Leaders</strong></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>Turn leaders against each other (by disinformation, deception).</td>
</tr>
<tr>
<td>The cause</td>
<td>Convince terrorist leaders that attacking the United States would undermine the cause. Raise the operational risks of attacking the United States.</td>
</tr>
<tr>
<td>Family, tribe, brotherhood</td>
<td>Convince state leaders to prevent rewards for families of terrorists. Convince state leaders to harass terrorist leaders and to punish them economically.</td>
</tr>
<tr>
<td><strong>Foot Soldiers</strong></td>
<td></td>
</tr>
<tr>
<td>The cause; the excitement</td>
<td>Raise the operational risks. With continued U.S. successes, demonstrate the futility of the cause.</td>
</tr>
<tr>
<td>Family, tribe, brotherhood</td>
<td>See above.</td>
</tr>
<tr>
<td><strong>External Suppliers and Facilitators (Financiers, Logisticians, etc.)</strong></td>
<td></td>
</tr>
<tr>
<td>The cause</td>
<td>Discredit the cause within Islam and society.</td>
</tr>
<tr>
<td>Wealth, power, life</td>
<td>Cause loss of wealth, loss of power, or loss of life through imprisonment, dishonor, or death.</td>
</tr>
<tr>
<td>Family, tribe, brotherhood</td>
<td>See above.</td>
</tr>
<tr>
<td><strong>Heads of Supportive States</strong></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>Conduct selected strikes and incursions (preemption). Impose military, political, and economic sanctions. Shun state supporters of terrorism. Credibly indicate that a military response to the use of weapons of mass destruction would be forever relentless and powerful.</td>
</tr>
<tr>
<td>Political goals</td>
<td>Convince state supporters of terrorism that attacks on the United States would undermine their goals. Provide other ways to reach goals.</td>
</tr>
<tr>
<td><strong>Supportive Population Segments</strong></td>
<td></td>
</tr>
<tr>
<td>Survival</td>
<td>Provide hope (peace process, economic aid, liberalization, etc.). Stimulate candid discussion about the likely blowback of biological attacks.</td>
</tr>
<tr>
<td>Bitterness, blame</td>
<td>Broaden the range of ideas and views discussed.</td>
</tr>
<tr>
<td>The cause</td>
<td>Remind populations of U.S. successes against al Qaeda and local successes against terrorists.</td>
</tr>
<tr>
<td><strong>Sources of Moral and Religious Support</strong></td>
<td></td>
</tr>
<tr>
<td>Power, status</td>
<td>Discredit religious leaders, warn them, monitor them, and shut off their funds.</td>
</tr>
<tr>
<td>Personal and family welfare</td>
<td>Cause imprisonment, dishonor, or death. Prevent benefits from accruing to families.</td>
</tr>
</tbody>
</table>

take action. Egypt and Pakistan have been increasingly cooperative, despite difficult internal challenges, and major arrests have recently been made in Pakistan. As of the time of our report, we judged Saudi Arabia to be a special case. On the one hand, the United States and Saudi Arabia have had a long and mutually beneficial strategic relationship. On the other hand, the spread of what is called in shorthand “Wahhabism,” long actively promoted by Saudi Arabia, has helped encourage intolerance that transmutes into religious extremism. The Saudi government (and governments throughout the Middle East and elsewhere) need to strongly discourage extremist and intolerant teachings and impede organizations that support terrorism.

Maintaining American Values at Home and Abroad. Core American values can be preserved in the war on terrorism—if not always, then in the main. America must be true to its own enduring values and must convince others that it is doing so. Broad international support is essential for success in this war.

On the home front, America must have—and play by—rules. America should consider more intrusive legal measures (such as preventive detention, military tribunals, and secret hearings)—but it should retain appropriate forms of due process and legal protections against abuses that would otherwise assuredly occur. America should also grant appropriate access to the press, which will enhance public faith in the government’s integrity and avoid dangerous rumors. Is all of this possible? In the late 20th century, every liberal democracy confronting terrorism (particularly in Europe) was obliged to modify its normal rules, yet it can be argued that each of the countries found it possible to do so without unduly sacrificing its values. This said, prosecutorial shortcuts taken out of necessity should not be allowed to become permanent or to be used as an excuse for broader limitations of civil liberties.

On the foreign front, the United States should continue to promote democracy. Many problems in the Middle East (and South Asia), including terrorism, are related to the lack of democratization. For years, the United States has failed to hold Middle Eastern states, in particular, to the same standards of democratization as it has held other nations. The time has come to begin doing so. The United States should increase its moral and fiscal support of United Nations and nongovernmental organizations that work to create the infrastructure of civil society in the Middle East. We believe that it is possible for the United States both to work effectively with current regional leaders and to encourage democratization more vigorously. And however difficult it continues to be to move toward resolution of the tragic Israeli-Palestinian conflict, negotiating some kind of resolution will continue to be of fundamental importance.

The United States should also exercise moral leadership in the military struggle against terrorism. Consistent with established American values in war, America can use force effectively but discriminately. America should recognize the need for reduced levels of proof before taking action—but it should avoid arrogant actions based on flimsy information. America should engage in information warfare to confuse enemy minds tactically—but it should focus as much as possible on the power of truth.

As we look at the situation a year after our original work on this topic, we note that the U.S. government has effectively adopted a systemic approach to counterterrorism. It has employed a broad range of strategies, distinguished among different elements of the al Qaeda system, and tailored tactics accordingly. It has proceeded along a broad front, rather than counting on finding a vulnerable core. The United States has also worked very closely with allies worldwide, with major payoffs. The United States has also sought—but with much less success so far—to engage in the war of ideas and to address underlying causes of terrorism, including the promulgation of intolerance and extremism, as well as the open sore that is the Israeli-Palestinian problem.

Related Reading


Since the 1980s, the herb ephedra has been purported to increase weight loss and to enhance athletic performance. In recent years, though, a string of athletes and other individuals who apparently had ingested the herb have collapsed and, in some cases, died. In the wake of the death of Baltimore Orioles pitcher Steve Bechler this spring, the voices of government officials and the public at large have grown increasingly louder in demanding proof of the safety and efficacy of the herb.

Unfortunately, scientific proof has become exceedingly difficult to attain. As a result of the 1994 Dietary Supplement Health and Education Act (DSHEA), substances that are classified as “dietary supplements” (including herbs such as ephedra) are not considered to be “drugs” and thus are not subject to the same rigorous regulatory standards as drugs are. According to the legislation, manufacturers of dietary supplements need not show evidence of the efficacy or safety of the

Preponderance of Evidence
Judging What to Do About Ephedra

By Paul G. Shekelle, Margaret Maglione, and Sally C. Morton

Paul Shekelle, a RAND and Veterans Affairs physician, is director of the Southern California Evidence-Based Practice Center. Margaret Maglione is a RAND policy analyst. Sally Morton holds the RAND Endowed Chair in Statistics.

Food and Drug Administration Commissioner Mark McClellan, foreground, with Health and Human Services Secretary Tommy Thompson, tells reporters on Feb. 28 that bottles of the popular dietary supplement ephedra should bear warning labels that the pills can cause heart attacks, strokes, or even death.
products prior to marketing them. Therefore, the usual regulatory pressure to produce clinical studies assessing efficacy and risk does not exist. Consequently, we could find only limited scientific data about ephedra.

However, we did find sufficient evidence of danger associated with ephedra, along with insufficient evidence of its efficacy for athletic performance, to compel the U.S. Department of Health and Human Services and the U.S. Food and Drug Administration (FDA) to take preliminary regulatory measures against the herb. On Feb. 28, 2003, the FDA proposed a strong new warning label on ephedra products, warned manufacturers against making unsubstantiated claims that the products could enhance athletic performance, and invited public comment on the risks associated with ephedra to support new restrictions on the products.

The FDA is now seeking public comment to determine whether dietary supplements containing ephedra present a “significant or unreasonable risk of illness or injury.” This is the standard that must be met under DSHEA before the FDA can take further regulatory action. This standard also reveals an indirect effect of the DSHEA law. By presuming that dietary supplements are safe and absolving manufacturers from proving that the supplements are safe, the law places the burden of proof on the FDA instead. The agency must somehow prove that a supplement is risky, even in the absence of clinical studies conducted by manufacturers.

Our study could not prove with scientific certainty that ephedra is unsafe. However, we compiled enough evidence to reach fairly confident conclusions. Our efforts could serve as an example of how policymakers...
We found some evidence of the benefits of ephedra and ephedrine for weight loss.

and researchers can help to keep the public safe despite the absence of incontrovertible scientific proof of danger.

Lengthy Deliberations

The active stimulant in ephedra is called ephedrine, which is found in over-the-counter drugs used to treat stuffy nose and asthma. The difference between the herb ephedra and the drug ephedrine is analogous to the difference between coffee beans and caffeine. Ephedra, known as “ma huang” among Chinese herbalists, is a shrub. Ephedrine is the active stimulant found in the shrub.

To determine if either substance could improve weight loss or enhance athletic performance, we searched the medical literature as well as other sources for published and unpublished clinical trials of the substances. We based our conclusions about efficacy on a detailed review of 52 trials of ephedra or ephedrine for weight loss or athletic performance.

Many of the 52 trials involved only small numbers of people, covered only short periods of time, or suffered from other limitations. For example, all of the trials for athletic performance involved just a couple of dozen young fit males, who were not representative of the general population. Even in aggregate, the 52 trials offered only weak evidence for assessing the relationship between rare adverse events and the use of ephedra or ephedrine.

To determine with greater confidence if it is safe to take ephedra or ephedrine, we analyzed nearly 18,000 case reports of adverse events. Consumers had contacted the FDA to provide the agency with 1,820 “adverse event reports,” the vast majority of which dealt with ephedra rather than ephedrine. The FDA shared these reports with us. We discovered 71 additional reports in the medical literature. The largest repository of reports came from a computer file of 15,951 cases reported to Metabolife, a San Diego–based maker of ephedra-containing dietary supplements.

Except in extraordinary circumstances, case reports cannot be considered to be conclusive evidence of a cause-and-effect relationship. However, case reports can be useful to establish the potential for a causal relationship.

Weight Loss

We found some evidence of the benefits of ephedra and ephedrine for weight loss. Dietary supplements containing ephedra alone, ephedrine alone, ephedra with herbs containing caffeine, or ephedrine plus caffeine promoted modest short-term weight loss of about two pounds per month more than among people taking a placebo. But none of the available studies followed participants for longer than six months. A study of 12 months is generally accepted as necessary to establish a drug’s value as a weight-loss aid.

Of the 52 clinical trials, 44 assessed ephedra, ephedrine, or ephedrine plus other compounds used for weight loss. Of these 44 trials, we excluded 18 because the duration of treatment was less than eight weeks. We excluded six other trials for a variety of reasons.

In the remaining 20 trials, we found comparisons made in six categories. We highlight the results below and in Figure 1.

- Ephedrine versus placebo. In five trials, the average weight loss for a person treated with ephedrine was 1.3 pounds per month more than the average weight loss for a person treated with a placebo.
- Ephedrine and caffeine versus placebo. In 12 trials, the average weight loss for a person treated with ephedrine and caffeine was 2.2 pounds per month more than the average weight loss for a person treated with a placebo.
- Ephedrine and caffeine versus ephedrine alone. In three trials, the average weight loss for a person treated with ephedrine and caffeine was 0.8 pounds per month more than the average weight loss for a person treated with ephedrine alone.
- Ephedrine and caffeine versus another active pharmaceutical for weight loss. In two trials, we found no statistically significant difference in weight loss. One trial compared the combination of ephedrine and caffeine to dexfenfluramine. The other trial compared the same combination to diethylpropion.
- Ephedra versus placebo. In a single trial, the average weight loss for a person treated with ephedra was 1.8 pounds per month more than the average weight loss for a person treated with a placebo.
- Ephedra with herbs containing caffeine versus placebo. In four trials, the average weight loss for a person treated with this ephedra mixture was 2.1 pounds per month more than the average weight loss for a person treated with a placebo.
In Figure 1, we show the average increase in weight loss in each case as a black square and give its value. We also include a vertical line that illustrates the 95-percent confidence interval around the average value. For example, the average additional weight loss for a person treated with ephedrine alone was 1.3 pounds per month. We are 95 percent confident that the true additional average weight loss would fall somewhere between 0.4 and 2.2 pounds per month.

**Athletic Performance**

We found no trials of ephedra on athletic performance—and thus no evidence that ephedra could enhance athletic performance. We found only minimal evidence that ephedrine could enhance athletic performance. Even here, there appeared to be no athletic benefit from ephedrine beyond an immediate boost.

There were eight trials of ephedrine on athletic performance. All but one included caffeine. Each trial involved different types of exercise and different outcome measures, so we analyzed each trial individually.

Six of the eight trials assessed the exercise capacity of small groups of healthy males. Each trial included 24 or fewer subjects. The trials concluded that neither caffeine nor ephedrine alone had significant effects on various parameters of exercise performance, such as oxygen consumption, time to exhaustion, or carbon dioxide production. However, the combination of ephedrine and caffeine consistently demonstrated a 20–30 percent increase in short-term performance.

One trial of strength training showed an improvement in muscle endurance—but only on the first of three repetitions. The remaining trial reported no statistically significant improvement in a battery of tests of physical function, including oxygen uptake, measures of endurance and power, reaction time, hand-eye coordination, speed, and self-perceived exertion.

**Safety**

The clinical trials of ephedra and ephedrine reported numerous adverse side effects. We grouped the symptoms into clinically similar categories, as follows:

- psychiatric symptoms: those described in the trials as euphoria, neurotic behavior, agitation, neuro-psychiatric symptoms, depressed mood, giddiness, irritability, or anxiety
- autonomic hyperactivity: those symptoms described as tremor, twitching, jitteriness, insomnia, difficulty sleeping, increased perspiration, or sweating
- palpitations: those symptoms described as palpitations, irregular heartbeat, loud heartbeat, heart pounding, or increased or stronger heartbeat
- hypertension: those symptoms described as hypertension, increased systolic blood pressure, or increased diastolic blood pressure
- upper gastrointestinal symptoms: those described as nausea, vomiting, abdominal pain, upset stomach, heartburn, or gastroesophageal reflux
- headache
- tachycardia: those symptoms described as tachycardia, or slightly elevated heart rate.

Figure 2 shows our estimates of the increased odds of suffering these adverse events when taking ephedra or ephedrine. Once again, we show each estimate as a black square and give its value. We also include a vertical line that illustrates the 95-percent confidence interval around the estimated value. For example, we estimate that the odds of a person suffering psychiatric symptoms were 3.6 times higher if the person took ephedra or ephedrine. We are 95 percent confident that the true value of the increased odds of suffering psy-
The additional evidence we gathered from the case reports raised even greater concerns about consumer safety. We screened the nearly 18,000 case reports and then reviewed in detail 284 reports of either death, heart attack, other cardiac symptoms, strokes, neurologic symptoms, seizures, or serious psychiatric symptoms.

We searched each of the 284 reports to assess whether ephedra or ephedrine was, indeed, a likely cause of the adverse event. We judged each case report by the following three criteria:

1. Documentation that an adverse event had occurred.
2. Either documentation that the subject had consumed ephedra or ephedrine within 24 hours prior to the adverse event or a toxicological examination revealing the presence of ephedrine or an associated product in the blood or urine. (For example, we did not require psychiatric symptoms to become manifest within 24 hours of using ephedra or ephedrine.)
3. Documentation that an adequate investigation had excluded other potential causes.

From the 284 reports of serious adverse events, we identified two deaths, three heart attacks, nine strokes, three seizures, and five psychiatric cases as sentinel events with prior ephedra consumption. We identified three deaths, two heart attacks, two strokes, one seizure, and three psychiatric cases as sentinel events with prior ephedrine consumption. About half of the sentinel events occurred in people aged 30 years or younger. We identified 43 additional cases as possible sentinel events with prior ephedra consumption and 8 additional cases as possible sentinel events with prior ephedrine consumption.

In aggregate, the case reports suggest a link between products containing either ephedra or ephedrine and catastrophic events, such as sudden death, heart attack, stroke, seizures, and serious psychiatric symptoms.

**Catastrophic effects of ephedra, including death, cannot be ruled out at a rate of less than one person per thousand.**
**The Verdicts**

Regarding weight loss, we found enough evidence to conclude that the short-term use of either ephedrine alone, ephedrine and caffeine combined, ephedra alone, or ephedra with herbs containing caffeine all promote weight loss in selected patient populations. However, all but three of the trials lasted for less than six months. Ideally, the trials should assess not only the results of a full year of treatment but also what happens after the treatment is discontinued.

Caffeine clearly adds additional efficacy to ephedrine in promoting weight loss. The effects of ephedrine and caffeine together are roughly equal to the effects of ephedra with or without herbs containing caffeine. Each results in about two pounds per month of weight loss over four months.

To put these pounds in context, though, competing FDA-approved weight loss drugs have been shown to be about equally as effective. The drugs sibutramine (Meridia) and orlistat (Xenical) have both resulted in average weight loss of 6–10 pounds over 6–12 months, and the drug phentermine (often used in combination with fenfluramine as "phen-fen") has resulted in average weight loss of 16 pounds over 9 months.

Regarding athletic performance, the few trials of ephedrine that we identified did not study the drug as used by the general population—that is, repeated use. Therefore, the effect of ephedra or ephedrine to enhance athletic performance over the long term is completely unknown.

Regarding safety, we conclude from the clinical trials that ephedrine and ephedra are associated with two to three times the odds of experiencing psychiatric symptoms, autonomic symptoms, upper gastrointestinal symptoms, and palpitations. It is not possible to separate out the effect that caffeine may contribute to these events.

We conclude from the case reports of ephedra and ephedrine that serious adverse events have occurred in young adults without other apparent causes. There may be a causal relationship between taking the substances and suffering rare serious adverse events. Catastrophic effects of ephedra, including death, cannot be ruled out at a rate of less than one person per thousand.

Our study has several limitations. As we note above, many of the clinical trials themselves had design limitations, and all of the weight loss trials were of short duration. In addition, the results of the weight loss trials may understate the dangers for the general population. These trials frequently involved medical screening to exclude people with preexisting conditions, such as heart disease, that could have predisposed the people to increased risks. It is unknown whether administering ephedra or ephedrine without such screening would increase the risks.

Despite these limitations, we found sufficient evidence to conclude that dietary supplements containing ephedra or ephedrine are associated with a modest increase in weight loss in the short term—but also an increase in a variety of serious health risks. We hope that our efforts in compiling medical evidence, even when it cannot be gleaned from clinical trials alone, can help others in the field find ways to reach similarly useful conclusions.

**Related Reading**


Bottles of Ripped Fuel Metabolic Enhancer, which contains ephedra, are shown in New York on June 18, 2002. A bottle of Ripped Fuel was found in the locker of Minnesota Vikings tackle Korey Stringer after he collapsed and died during a training camp practice on Aug. 1, 2001.
Innovative attempts to improve education have flourished over the past decade. Unfortunately, many of the attempts have failed to produce all the desired results. Nevertheless, our assessments of two of the attempts have yielded important, if contrasting, lessons for policymakers.

In one attempt, chief executives from some of America’s most successful businesses launched a campaign to redesign public schools nationwide. In the other attempt, community groups working with the United Way launched a campaign to improve early care and education for low-income children in and around Pittsburgh, Penn.

Both campaigns met with a mix of success and failure. In both cases, flawed “theories of action” led to some wasted effort. In general, the business-driven campaign had a poor community plan, whereas the community-driven campaign had a poor business plan. Mistaken assumptions hobbled both efforts from the outset, but both of the experiences are enormously instructive for future large-scale educational reform initiatives.

New American Schools

About a decade ago, a private nonprofit corporation called New American Schools (NAS) set out to reverse the perceived lagging performance of American students and the lackluster results of school reform efforts. Funded largely by private-sector donations, NAS was formed in 1991 in conjunction with the America 2000 initiative of former President Bush.

The mission of NAS was to help schools and districts raise student achievement levels by implementing “whole-school designs.” NAS founders believed that, in the past, many school reforms were “programmatic,” or focused only on a particular set of individuals, a particular subject, or a particular grade level. This approach of adopting multiple and unconnected programmatic reforms allegedly resulted in a fragmented curriculum, a balkanized school organization, and low performance by students. Instead, the NAS founders believed that high-quality schools possessed a “unifying design” that integrated school practices into a coherent and mutually reinforcing set of effective approaches to teaching and learning for the entire school. Under these conditions, the staff could presumably function to the best of their abilities.

This “theory of action” was compelling. The critical assumption, of course, was that the coherent, focused, and sustained implementation of key design features—such as professional development, curriculum and instructional materials, content and performance standards, regular assessments, organization and governance, and parent and community involve-
ment—would improve school and classroom environments and thereby raise student achievement.

During the first year of the initiative, NAS held a competition and made awards to 11 design teams that were then commissioned to work with schools. At the schools, implementation consisted of putting into practice the core features of a whole-school design as developed by the NAS-affiliated design teams.

These efforts were expected to have measurable effects on student achievement. The private-sector sponsors wanted results in the form of many schools adopting the designs and showing improved student performance within five years. After five years, NAS was supposed to go out of business. NAS planned for a development phase (1992–1993), a demonstration phase (1993–1995), and a scale-up phase (1995–1998). Since then, NAS has continued to exist, but in a very changed form and with a changed purpose.

During the different phases, the board enjoined NAS staff to delete design teams that could not deliver on the promises of their designs, that had a limited potential market, or that could not show an ability to become financially independent from NAS. NAS’s board made clear that it was not interested in supporting a group of financially dependent design teams. The board insisted that the teams would soon move to a fee-for-service arrangement.

Over time, several teams were dropped from the NAS initiative. For example, NAS refused to support designs that were not transferable to schools across the country. NAS also refused to support design teams that were led by a central district office or a particular state. Only teams that were external to a local governance structure and that were serious about scale-up outside a “home” district or locality were acceptable.

During the scale-up phase (in which only seven teams remained), NAS used a district-level strategy by partnering with Cincinnati, Ohio; Dade County, Fla.; Kentucky; Maryland; Memphis, Tenn.; Philadelphia, Penn.; Pittsburgh, Penn.; San Antonio, Tex.; San Diego, Calif.; and three districts in Washington State. By 1995, about 185 schools in these jurisdictions had implemented NAS designs, while NAS design teams had spread to over 550 schools nationwide. By 1999, the teams were partnering with over 1,000 schools across the country.

Near the end of its originally planned life, NAS argued successfully for a new federal program to fund schools to adopt NAS-like designs. In 1997, the Comprehensive School Reform Demonstration Program was created to provide schools with funding to implement designs similar to those created and developed by NAS. Federal funding allowed the NAS designs to spread to over 4,000 schools by 2001.

To prove that its approach was efficacious, NAS asked RAND to assess (1) whether the whole-school designs could be developed, (2) whether they could be implemented, and (3) whether they could lead to improved student outcomes. We summarize the results of the RAND studies below.

Developing the Designs. While the designs were developing, they kept adapting. Positive adaptations included an increase in the amount of “design-based...
Clear communication and strong assistance by design teams fostered stronger teacher support.

assistance” offered to schools by the NAS design teams, continued refinement of curricular units, and the development of protocols to help schools choose their designs. Other adaptations became problematic. Adapting to district and school policies led some designs to accept unaligned and incoherent mixes of standards, assessments, curricula, instruction, and professional development.

Because of the latter adaptations, a major component of NAS’s theory of action—a coherent, unifying design—was often missing or was constantly in the process of being revised. It cannot be emphasized enough that during the entire time of the RAND studies, designs were still in a state of development.

Implementing the Designs. During the scale-up phase, NAS and the design teams partnered with schools and districts that were characterized by a host of problems related to poverty, low achievement, and other challenges. Two years into implementation, only about half of the sample schools in our case studies were implementing the designs at a level consistent with expectations. The other half were below this level. All schools reported additional barriers to further implementation. Four years into the effort, schools on average reported only modest levels of implementation.

Many factors in the initial process of selecting a design influenced its final outcome. Clear communication and strong assistance by design teams fostered stronger teacher support, which led to higher levels of implementation. The schools that reported a well-informed choice process reported higher levels of implementation than those that reported being forced to accept a design or not understanding the nature of a design. Principals also played a crucial role in ensuring a sound selection process and in gaining the initial buy-in of teachers.

School characteristics were associated with implementation. A teacher survey indicated that strong principal leadership was the single most important indicator correlated with the schoolwide level of implementation and that such leadership was also associated with reducing the variance in implementation among teachers within a school. The degree of implementation also correlated with teacher perceptions of the students and of their readiness to learn. Implementation was also higher in elementary schools than in secondary schools and in smaller schools than in larger schools.

The district context was important as well. Implementation was higher in districts that were more supportive of NAS designs and that had stable leadership, no budget crises, a coherent program of reform, resources dedicated to the NAS effort, significant school-
level autonomy, and a trusting relationship among three groups: school staff, district staff, and union staff.

Meanwhile, many other reforms took place in the schools concurrently, often causing teacher overload and reducing their capacity to implement the NAS designs. A classroom study in San Antonio revealed that the adoption of multiple reforms at the district level easily overwhelmed teachers and their efforts at this particular reform. Most important, the high-stakes state testing regime, which encouraged a focus on basic skills, resulted in a district adoption of curricular programs in addition to the NAS designs. In fact, the curricular programs associated with the state testing regime conflicted with the NAS designs and resulted in lower levels of implementation.

Lack of funding was the single most important reason cited by most schools in the decision to drop a design. Significant unfunded costs were borne by teachers in terms of their time and effort involved.

**Improving Student Outcomes.** Of the 163 NAS schools for which we had data to compare performance relative to their districts, 81 schools (50 percent) made gains relative to the district in mathematics, and 77 schools (47 percent) made gains in reading (see Figure 1). Among the jurisdictions with 10 or more NAS schools, Memphis and Kentucky appeared to improve the most in mathematics, while Cincinnati and Washington State appeared to do best in reading. Better and longer-term data are needed to make conclusive judgments about the effects of the designs on student performance.

Our detailed classroom study of San Antonio allowed us to examine the effects of variations in instructional conditions. We found that strong principal leadership, as reported by teachers, had significant positive effects on student test scores in reading and mathematics. We found that the instructional conditions promoted by reforms such as NAS—including teacher-reported collaboration, professional development, and revised instructional practices—were not related to student achievement net of other student and classroom conditions. We also found that early implementation of NAS designs in a high-poverty district within a high-stakes testing and accountability system did not result in significant effects on student achievement.

**Lessons Learned.** Our findings provide mixed evidence to support the NAS theory of action. One of the primary lessons is that the designs, by themselves, cannot transform schools. Beyond the designs, the schools also need significant amounts of professional development, technical assistance, and materials geared to the implementation of the designs.

We conclude that the NAS theory of action was largely underdeveloped. The causal chain of events leading to strong implementation and improved student outcomes proved to be far more complex than originally imagined by NAS. The chain of events also remained largely outside the control and influence of NAS. Exogenous variables included school and district leadership; demographic, financial, and other prevailing conditions; the simultaneous adoption of multiple reforms; and the sometimes conflicting implementation of high-stakes testing regimes.

Nonetheless, the experiment has produced the following important lessons for future large-scale educational reform efforts:

**Designs, by themselves, cannot transform schools.**

**Externally developed interventions cannot “break the mold” and yet still be implemented in the existing contexts of schools and districts.** The evidence suggests that schools were not by and large fertile ground for “break the mold” ideas, often because of a lack of capacity or because of local, state, or district regulations. The schools did not have a ready place for the designs. Instead, the designs had to adapt to school conditions or simply not be implemented.
The high-stakes testing regime precluded the adoption of rich and varied curricula. External interventions need to address systemic issues that can hinder implementation. Systemic issues include the lack of teacher capacity, especially in terms of time and subject area expertise; the lack of strong principal leadership; and a district infrastructure at odds with the needs of a design. A rush to scale up when interventions are not completely developed weakens results. Many of the problems associated with the NAS scale-up phase were the result of pushing toward full-scale adoption before the designs were developed and before the design teams had created the capacity to support more schools. A key component of successful implementation is consistent, clear, and frequent communication and assistance between design teams and schools, particularly teachers. A strong, trusting relationship between a school and an external agent is a prerequisite for strong implementation of complex interventions that require significant changes in behavior. If external players expect teachers to change their behavior significantly, then the external players need to invest considerable time and effort in building relationships with teachers. Monitoring site progress is necessary if developers are to succeed and to improve their designs over time. Unless systems for tracking progress and for understanding school-level concerns are created and used for improving the intervention, the effort cannot succeed over the long term. The typical outcome measures used in public accountability systems provide a very limited measure of student and school performance. In the sample we studied, the high-stakes testing regime precluded the adoption of the rich and varied curricula developed by design teams—curricula that could have challenged students and motivated them toward more in-depth learning experiences. The overwhelming emphasis now given to the scores on state- or district-mandated tests does not bode well for many innovative reform efforts.

Early Childhood Initiative
The Early Childhood Initiative (ECI) aimed to deliver high-quality early care and education services to low-income children from birth through age five in Pittsburgh and the surrounding communities of Allegheny County. ECI planners hoped to prepare the children for kindergarten, to promote their long-term educational attainment, and to give them the early tools needed to grow into productive, successful members of society.

From 1996 through 2000, ECI operated under the auspices of the United Way of Allegheny County. By April 2001, the program had been scaled down, converted to a demonstration program, and placed under the management of the University of Pittsburgh. The goals of the program had been quite ambitious. It aimed (1) to provide high-quality services, (2) to do so on a large scale (to serve 7,600 children in 80 neighborhoods within five years), (3) to do so inexpensively (at a cost of $4,000 to $5,000 per child per year), (4) to use a community-driven approach, and (5) to achieve sustainability through a commitment of state funding.

ECI succeeded in one important respect: The program generated high-quality services for hundreds of children in several communities. However, ECI failed in many other respects, despite the good intentions of everyone involved and the support of a wide array of community leaders. Four years after its launch, ECI was far short of its enrollment targets. The cost per child was much higher than anticipated. And the efforts to secure long-term state funding had failed.

There are several policy lessons. Given the goals of ECI, success required that it have a clear sense of market realities, a well-designed theory of action, an effective strategy to induce public funding, and a coherent organizational structure. Weaknesses in each of these areas undermined ECI’s success in meeting its goals that pertained to scale, cost, community, and sustainability.

A Clear Sense of Market Realities. ECI’s business plan made assumptions about the population to be served, the services to be delivered, and the participation of existing child-care providers. Many of the assumptions proved to be incorrect because the planners had paid insufficient attention to the demand for services, the supply of services, and the incentives to use services.

The plan failed to anticipate that parents and neighborhood agencies would gravitate toward the highest-cost service provided by ECI. The services ranged from part-day, Head Start–like enrichment and literacy programs to full-day, center-based care and education. The plan assumed that 71 percent of children would be served in low-cost, part-day programs. In fact, virtually all the children were served in full-day programs, most in new child-care centers created by ECI.
ECI underestimated the demand for the full-time programs in part because it had underestimated the proportion of eligible children whose mothers were in the workforce. In addition, a parental preference for full-day care is unsurprising if parents are given a choice of full-day or part-day care and both options are largely or entirely subsidized. Large numbers of parents—whether of low or high income, employed or not—are likely to prefer more hours of child care and education if they are offered at little or no additional cost. This is essentially the choice that ECI offered.

As a result, ECI delivered more services per child than expected and thus cost more per child than expected. In 1999, the average cost per child per year was $13,612. That sum is not dramatically different from the cost of other, high-quality services of this kind, but it is three times as high as the cost anticipated in the ECI business plan.

The plan also assumed that many existing child-care providers would be used; instead, many were left out, for several reasons. First, many existing providers operated at so low a level of quality that they were deemed incapable of providing high-quality services. Second, some community groups refused to include the existing providers. Third, some providers refused to participate, either because they considered the quality standards and monitoring process too intrusive or because they were informal, unregistered providers not wishing to become part of the formal child-care system. Therefore, ECI incurred the high costs of creating new centers.

The plan further assumed that centers that were not fully enrolled could serve children for the same cost per child as centers that were fully enrolled. This assumption might have made sense for preexisting centers, which ECI had intended to reimburse on a per-child basis. But the great majority of ECI children attended newly established child-care centers, where the operating costs were largely fixed regardless of the number of children attending. ECI centers almost never reached 100 percent of capacity. In 1999, the average enrollment was 73 percent of capacity. As a result, operating cost per child was higher than expected.

A Well-Designed Theory of Action. At the time of ECI’s inception, no models existed of high-quality early childhood services delivered on a large scale through grassroots, neighborhood control. ECI therefore needed to develop its own theory of action to explain how the initiative would work. According to its theory of action, ECI would have a central administration housed within the United Way. The administration would fund, supervise, and monitor (with stringent quality standards) the lead agencies in each community (see Figure 2). Each lead agency would in turn fund and supervise the participating providers. The providers would then serve the children and their families. The central administration would also provide technical assistance to the lead agencies, which would in turn provide such assistance to the community providers.

This theory of action proved to be cumbersome and problematic. The ultimate goal of delivering good service pertained mostly to the community providers and the families, but the theory of action pertained mostly to the central administration and the neighborhood lead agencies. The theory of action put several layers of organization between the funding sources and the primary intended beneficiaries (i.e., the children and parents) and produced a number of implementation problems.

Another consequence of the multilayered theory of action was that each layer added to the administrative cost. The theory led to the imposition of substantial top-down costs.
requirements and created additional administrative structures in each neighborhood, virtually guaranteeing that the administrative costs would be high.

Figure 3 shows how each of the mistaken assumptions of the business plan contributed to the cost overruns. The first bar shows the original budget. The second bar shows that the shift to full-day care dramatically increased both the operating and the capital costs. The third bar adds the additional costs of operating new child-care centers. These two factors alone explain 65 percent of the cost overruns. The remaining 35 percent can be attributed to higher central administrative costs, neighborhood-level costs, and other capital costs.

An Effective Strategy to Induce Public Funding. ECI planned to spearhead a lobbying effort to persuade the state of Pennsylvania to fund the initiative at the end of its five-year start-up period, when the initial infusion of money from foundations and private donors would be exhausted. ECI failed in this effort, both because of the obstacles involved and because of the strategy pursued.

The obstacles were numerous. To begin with, the benefits of early childhood programs are diffuse, whereas the costs are concentrated. Even large social benefits might not be sufficient to persuade a state legislature to fund a program when many of the benefits will not accrue to the state's treasury. Most of the benefits also accrue over the long term, whereas the costs are borne immediately. The amount of funding requested by Allegheny County would inevitably raise issues of regional fairness among Pennsylvania political leaders. And voters, like policymakers, remain ambivalent about public funding of early childhood programs. Some voters and policymakers feel strongly that child care is a private responsibility of parents rather than a public responsibility of the state.

Independent of these obstacles, ECI planners failed to ensure that state policymakers had a full, substantive, and early role in the initiative's design. Even more self-defeating, the goal of ECI conflicted with the goal of Pennsylvania's system of child-care subsidies for low-income families. ECI's goal was to provide high-quality early education to low-income children—regardless of whether their parents were working. Pennsylvania's goal for child-care subsidies, in contrast, was to prod parents on public assistance into the workforce. Even if ECI's planners disagreed with Pennsylvania's goal, the planners needed to recognize that a direct conflict with the state goal would seriously undermine the likelihood of state support for ECI.

This external conflict became a serious problem for ECI as it became increasingly dependent on the state's subsidy system. The United Way tried to make ECI compatible with that system, but such efforts created major internal conflicts and undermined ECI's support in the neighborhoods.

A Coherent Organizational Structure. ECI lacked an independent board with the authority to resolve conflicts and make key decisions. A proliferation of volunteer oversight committees added to the administrative confusion and further diffused the authority.

The complex leadership structure led to several problems. Both the community planning process and the response to changing conditions were slow, because the plans and responses had to be reviewed by several layers of committees. Confusion over who had the authority to make what decisions allowed disagreements to escalate into full-blown, unresolved power struggles among managers. Communication with funding sources and business leaders broke down because managers disagreed over what they should be communicating. Neighborhood agencies and providers received mixed signals about the rules and procedures. Predictably, support for ECI declined among funding sources, business leaders, and neighborhood groups.

The benefits of early childhood programs are diffuse, whereas the costs are concentrated.
At the same time that it had established a hierarchical bureaucracy, ECI had aimed to permit neighborhoods to direct the local programs. But devolution of authority to the community level requires a trade-off: Neighborhood-led programs might be more robust and effective than those imposed from without, but implementation is not likely to proceed rapidly. ECI planners failed to appreciate how much time neighborhood groups would need to mobilize, to assess their needs, to find space for child-care centers, to develop proposals, and to establish programs.

ECI’s business plan failed to acknowledge the extent to which quality control and community control might be in tension. ECI’s insistence on its own definition of quality placed constraints on the freedom to be exercised by the communities. Some neighborhoods were disappointed when they discovered that their dreams were not always consistent with the vision of ECI.

Lessons Learned. The foregoing discussion suggests at least four lessons for future large-scale reform initiatives:

- Careful consideration of demand, supply, and the likely responses to incentives is essential to anticipating unintended consequences.
- A thoughtful and realistic theory of action promotes an initiative’s goals more effectively.
- Bold visions require hardheaded plans that acknowledge the political and policy realities and that include all relevant stakeholders early in the planning process.
- An ambitious, large-scale initiative should have an independent board and a clear administrative structure that promotes strong leadership.

Whither NAS and ECI?

Across the country, the New American Schools would have had greater success if the plan had been more flexible with respect to the prevailing conditions within schools and districts, more forgiving with respect to the time lines, and more attentive to the concerns of teachers.

In Pittsburgh, the Early Childhood Initiative would have had greater success if the community plan had incorporated a clear sense of market realities, a well-designed theory of action, a better strategy to induce public funding, and a stronger organizational structure.

Because of poor planning, the expectations regarding these reforms were not likely to be met. But the planning can be improved in the future. Both of these reforms—one targeting entire schools, the other targeting children from birth to age five—have succeeded partially despite missteps along the way. Federal and state policymakers should not abandon what could be promising vehicles for reform before giving them a fair chance to prove themselves.

Quality control and community control might be in tension.

Related Reading


While talk about weapons of mass destruction dominates our attention today, "weapons of mass creation"—the potential to change the future of the human race through genetic manipulation—may, in some ways, be even more threatening. Just as the splitting of the atom unleashed great possibilities and terrible destructive powers, so the unraveling of the double helix offers us similarly daunting choices.

Most of the current debate focuses on cloning for reproduction, while genetic manipulation—sometimes called reprogenetics—is being actively pursued without much debate. The debate is limited because, for now, reprogenetics is focusing on curing diseases such as sickle cell anemia and Huntington’s and offering regenerative medicine to Parkinson’s victims.

But successes in eliminating or curing diseases will, down the road, lead to efforts to alter genes to enhance human traits. Many parents, desiring the best for their children, will want to use genetic manipulation to make their unborn children smarter, more creative, more attractive, or more athletic. Altering the random genetic combinations that make us who we are could be disastrous for society as a whole.

Already many couples in Asia believe that having a son is a justifiable family goal, in part because sons will do more to support the parents in old age. Some achieve this goal by using genetic testing and selective abortions to produce boys. The result, according to the Global Health Council, is significant gender imbalances in many nations: China has 118 boys per 100 girls under age 5, South Korea has 117 to 100, and Taiwan has 110 to 100. Growing numbers of what demographer Nicholas Eberstadt calls “unmarriageable males” will have little hope of finding wives, leading to a range of social problems.

Now fast-forward several decades. Imagine your neighbors genetically enhancing their children’s physical attractiveness and brainpower. Wouldn’t it be hard to resist the opportunity to offer your children the same competitive advantages—or at least a level playing field?

The option to alter the genes that enhance desirable characteristics will almost surely be available, at least initially, only to the wealthy, creating what Lee Silver, a Princeton University biologist and genetic expert, calls the “GenRich.” The wealthy will use technology to ensure their children have significantly more advantages than the random mix of the gene pool would provide.

The fallout from this “weapon of mass creation” will be a vastly more unequal world. The very founding principle of our society—that we are all created equal—will become a quaint notion. The temptation of the genetically enhanced to anoint themselves as leaders and protectors of their “less equal” fellow citizens could prove to be overwhelming.

And if scientists in one nation are genetically altering unborn children, wouldn’t it be hard for other nations to resist joining in a “genetic arms race” to develop a new generation better able to compete in the global economy—or in war?

Although reprogenetics is still in the early stages of research, scientists report few obstacles to eventual success. We are standing at the gateway of a brave new world of genetic manipulation that could, in this century, create what Francis Fukuyama calls our “post-human” future.

Before we step through the gate, we should look ahead and determine the consequences of the paths we could take and ask who gets to decide. Genetic alteration, even with altruistic aims, needs careful oversight and governance with the power of law—and not just by watchful scientists or government councils. Citizen councils that have broad representation and cross-national linkages—such as European councils on genetically modified foods—should be actively involved, and it is not too soon to begin. Anything less may well bring about a world where a post-human race dominates the rest. Which side will your progeny be on?
The Global Threat of New and Reemerging Infectious Diseases
Reconciling U.S. National Security and Public Health Policy
Jennifer Brower, Peter Chalk

The spread of infectious diseases—including some once thought to be eradicated—is a growing menace garnering increased attention and concern. The authors review the changing nature of security, focusing on two case studies—HIV/AIDS in South Africa, as an example of pervasive and insidious disease as a threat to state security; and the U.S. capacity to respond to infectious diseases—as well as policies, institutions, and procedures that can mitigate the threat.

2003 • 163 pp. • ISBN 0-8330-3293-3 • $20.00 • paper

Deterrence & Influence in Counterterrorism
A Component in the War on al Qaeda
Paul K. Davis, Brian Michael Jenkins

This overview discusses crosscutting strategies that can be used to deter or to influence the al Qaeda system, including supporters. The authors review broad elements of strategy, countermeasures ranging from co-optation to crushing attacks, and issues such as mass-casualty weapons and political warfare. Three appendixes place counterterrorism strategy in the framework of cold war deterrence theory, complex adaptive systems, and effects-based planning.

2002 • 105 pp. • ISBN 0-8330-3286-0 • $20.00 • paper

Unweaving the Web
Deception and Adaptation in Future Urban Operations
Scott Gerwehr, Russell W. Glenn

Deception is a powerful yet understudied instrument of war. Reviewing deception in the animal kingdom—where ruses of a near-infinite variety are applied to offense, defense, and intelligence gathering—the authors delve into the theory of deception to reveal new avenues of exploration applicable at every level of war.

2003 • 98 pp. • color plates • ISBN 0-8330-3159-7 • $24.00 • paper

New Challenges, New Tools for Defense Decisionmaking
Stuart E. Johnson, Martin C. Libicki, Gregory F. Treverton (editors)

It is still easy to underestimate how much the collapse of the Soviet Union and the end of the cold war transformed the task of American foreign and defense policy. The terrorist attacks of Sept. 11, 2001, launched a new era, the shape and dimensions of which are yet to be fully understood. This volume addresses the challenges of this changed world, the difficulties for defense planning that these challenges engender, and new analytic techniques for framing these complex problems.

2003 • 414 pp. • ISBN 0-8330-3292-5 • $40.00 • cloth
ISBN 0-8330-3289-5 • $30.00 • paper

Terrorism & Development
Using Social and Economic Development to Inhibit a Resurgence of Terrorism
Kim Cragin, Peter Chalk

The authors review social and economic development policies enacted in three countries—Israel, the Philippines, and the United Kingdom—to inhibit a resurgence of terrorist violence within their territories. The varied results provide lessons for future development efforts, especially those designed for purposes of conflict resolution and the prevention of terrorist violence.

2003 • 66 pp. • ISBN 0-8330-3308-5 • $18.00 • paper

Turkish Foreign Policy in an Age of Uncertainty
F. Stephen Larrabee, Ian O. Lesser

The authors, longtime observers of Turkey and the Mediterranean region, describe the challenges and opportunities facing Turkey in the international environment during a time of extraordinary flux. Special emphasis is given to the strategic and security issues facing Turkey, including a number of issues posed by the terrorist attacks of Sept. 11, 2001, and the subsequent international response.

2003 • 239 pp. • ISBN 0-8330-3281-X • $24.00 • paper