

COALITION FOR EVIDENCE-BASED POLICY

A Project Sponsored by

THE COUNCIL FOR
EXCELLENCE
IN GOVERNMENT

Bringing Evidence-Driven Progress To Crime and Substance-Abuse Policy: A Recommended Federal Strategy

December 2003

*The development of this report is funded by a grant
from the Robert Wood Johnson Foundation
and the Jerry Lee Foundation.*

Board of Advisors

Robert Boruch
University of Pennsylvania

Jonathan Crane
Progressive Policy Institute

David Ellwood
Harvard University

Judith Gueron
Manpower Demonstration
Research Corporation

Ron Haskins
Brookings Institution

Robert Hoyt
Jennison Associates

David Kessler
University of California, San
Francisco

Jerry Lee
WBEB 101.1 FM Philadelphia

Diane Ravitch
New York University

Laurie Robinson
University of Pennsylvania

Isabel Sawhill
Brookings Institution

Martin Seligman
University of Pennsylvania

Robert Slavin
Johns Hopkins University

Robert Solow
Massachusetts Institute of
Technology

Nicholas Zill
Westat, Inc.

Executive Director

Jon Baron
jbaron@excelgov.org

1301 K Street, NW
Suite 450 West
Washington, DC 20005
202-728-0418
FAX 202-728-0422

www.excelgov.org/evidence

Executive Summary

In January 2003, the Coalition for Evidence-Based Policy and the Justice Department's Office of Justice Programs launched a joint initiative to advance evidence-based crime and substance-abuse policy – an area where progress has been thwarted by government programs implemented with little regard to rigorous evidence. Top officials from the White House Office of National Drug Control Policy, Department of Health and Human Services, and Department of Education are also participating in this initiative. Its purpose is to explore how the federal government can most effectively use its resources to advance the development and effective use of rigorous evidence on what works in crime and substance-abuse policy. This report sets out specific recommendations for consideration by the participating agency officials and the broader policy community, including Congress. While the report is based on extensive input from federal officials and others, its conclusions and recommendations are those of the Coalition.

The Problem: Crime and substance abuse inflict profound and continuing costs on life and health in America.

Specifically, substance abuse – defined broadly to include excess alcohol consumption, drug abuse, and smoking – accounts for approximately 25 percent of U.S. deaths each year. In addition, Americans are victims of over 5 million violent crimes each year, including 15,000 homicides. Government data show that the United States has made no significant progress in decreasing substance abuse over the past decade, either among youth or in the general population. The United States has made progress in reducing violent crime, but evidence suggests that evidence-based crime policy could yield even greater reductions in the future.

Progress is often thwarted by government programs and strategies (“interventions”) that are not based on rigorous evidence.

For example, the nation's most widely-used school-based drug use prevention program – Drug Abuse Resistance Education (DARE), which receives \$200 million annually in public support and operates in 75% of U.S. school districts – has been shown in randomized controlled trials to have little or no effect on drug use by program participants. Finally, after 17 years of operation, DARE announced in early 2001 that it would change its curriculum in response to these studies. As another example, the nation's juvenile justice system typically places severely delinquent adolescents in group homes or other congregate care settings – a practice that randomized trials have shown may actually *increase* criminal behavior.

Most crime/substance-abuse interventions, however, have never been rigorously evaluated. For example, a 1997 Justice Department report which reviewed over 500 impact evaluations in the area of crime prevention concluded that: "The effectiveness of most crime prevention strategies will remain unknown until the nation invests more in evaluating them. . . . By scientific standards, there are very few programs of proven effectiveness."

We propose a major federal strategy to:

- (i) Build the knowledge base of crime and substance-abuse interventions shown effective in randomized trials** – not just in demonstration projects but when replicated in community settings; and
- (ii) Spur the widespread use of such evidence-backed interventions by recipients of federal crime/substance-abuse funding.**

The specific recommendations comprising this strategy are outlined later in this executive summary.

This strategy may offer a key to bringing sustained, evidence-driven progress – for the first time – to U.S. crime and substance-abuse policy.

Randomized trials have identified a few social interventions that are highly effective in addressing the problems of crime and substance abuse. Although rare, their very existence suggests that a concerted government effort to build the knowledge base of these evidence-backed interventions, and spur their widespread use, could fundamentally improve the effectiveness of federal crime/substance-abuse policy. Illustrative examples of these interventions include:

- **Nurse-Family Partnership** – a nurse visitation program for low-income women during pregnancy and children’s infancy (at 15-year follow-up, reduces children’s arrests, convictions, number of sexual partners, and alcohol use by 50-80%, compared to controls).
- **Life Skills Training** – a substance-abuse program for junior high students that teaches social and self-management skills, techniques for resisting peer pressure, and consequences of drug use (reduces smoking by 20% and serious levels of substance abuse by 30-50% by end of high school, compared to controls).
- **Prison Therapeutic Community** – a program that creates a separate community within a prison for inmates with drug problems who are scheduled for release, provides counseling/instruction for up to one year after release, and is staffed by highly committed role models of recovering substance abusers (at two-year post-release, reduces reincarceration by 35%, compared to controls).

Such examples of demonstrated effectiveness are rare because randomized trials are relatively uncommon in crime/substance-abuse policy. Meanwhile, careful empirical investigations show that the study designs that are commonly used (including pre-post designs and most comparison-group designs) often produce erroneous conclusions and can lead to practices that are ineffective or harmful. *Well-matched* comparison-group designs can produce valuable knowledge, but studies show that their results, too, need to be confirmed in randomized trials wherever possible.

Our main recommendations are:

1. **That the federal agencies develop a concise, uniform, user-friendly set of principles on what constitutes “rigorous evidence” of an intervention’s effectiveness.** Based on the strong empirical evidence, noted above, regarding the limitations of nonrandomized studies, the principles should recognize well-designed randomized trials as the basis for “strong” evidence of effectiveness (and well-matched comparison-group studies as providing “potential” evidence of effectiveness). These principles would provide – for the first time – a clear, authoritative source of federal guidance that could greatly accelerate both the development and use of evidence-backed interventions.
2. **That the agencies – both individually and together – launch a major strategy to build the knowledge base of evidence-backed crime/substance-abuse interventions.** Given the cost of large randomized trials, the agencies should invest *strategically* in such trials and related research. Specifically, each agency should (i) identify high-priority areas for building this knowledge base, drawing, for example, on what previous research suggests will be the areas of greatest payoff; (ii) invest initially in small randomized trials and well-matched comparison-group studies to create a pool of promising candidate interventions for large trials; and (iii) invest in large, well-designed trials only where previous evidence suggests likely success, and where feasible.

In addition, the agencies with large research budgets and agencies with large program budgets should undertake *coordinated* initiatives, that fund researchers and state/local agencies to join forces to carry out randomized trials of new interventions in community settings.

3. **That each agency focus its discretionary funds for research/evaluation, to the maximum extent practicable, on the above strategy to build the knowledge base of evidence-backed interventions.** Also, agency grant programs should give applicants major incentives to focus *their* discretionary funds on this strategy (e.g., additional funding, competitive priority in the proposal selection process, and waivers of certain statutory/regulatory requirements for undertaking randomized trials).
4. **That each agency establish, or contribute to, a “What Works” web site that provides authoritative, user-friendly information to practitioners on evidence-backed interventions.** As part of this effort, the Justice Department should establish a What Works web site on evidence-backed crime interventions. Such web sites could play a vital certifying function, identifying evidence-backed interventions for practitioners who would not otherwise have the expertise or resources to review the evidence themselves. Each What Works site should indicate which of its listed interventions are supported by “strong” evidence under the uniform federal principles developed per recommendation 1.
5. **That agency crime/substance-abuse grant programs, where appropriate, require applicants to provide a concrete strategy for implementation of evidence-backed interventions with fidelity.**
6. **That each agency undertake a major effort to educate the policy and grantee communities on the value of these evidence-based reforms, and provide technical assistance to facilitate their implementation.**

Conclusion:

The recommendations are all designed to be implemented in the near term, and, if coordinated among the federal agencies, can be implemented within the agencies’ existing statutory authority and funding levels. But they will require sustained attention and commitment by the agency leadership and staff. This effort, we believe, would spark cumulative, rapid progress in addressing a problem that is responsible for one-quarter of all U.S. deaths each year, and many damaged lives.

Bringing Evidence-Driven Progress To Crime and Substance-Abuse Policy: A Recommended Federal Strategy

I. Background: This is the final report of an initiative of the Coalition for Evidence-Based Policy and senior federal officials to advance evidence-based crime/substance-abuse policy.

In January 2003, the Coalition for Evidence-Based Policy and the Justice Department's Office of Justice Programs launched a collaborative initiative to advance evidence-based crime and substance-abuse policy. Top officials from several federal agencies are participating as principals in this initiative, including:

Deborah Daniels, Assistant Attorney General for the Office of Justice Programs, Department of Justice (DOJ) – the lead agency in this initiative
Andrea Barthwell, Deputy Director for Demand Reduction, White House Office of National Drug Control Policy (ONDCP)
Wilson Compton, Director, Division of Epidemiology, Services and Prevention Research, National Institute on Drug Abuse (NIDA)
Charles Curie, Administrator, Substance Abuse and Mental Health Services Administration (SAMHSA)
Wade Horn, Assistant Secretary for the Administration for Children and Families, Department of Health and Human Services (HHS)
Thomas Insel, Director, National Institute of Mental Health (NIMH)
Bill Modzeleski, Associate Deputy Under Secretary, Office of Safe and Drug Free Schools, Department of Education (ED)
Nora Volkow, Director, National Institute on Drug Abuse
Russ Whitehurst, Director, Institute of Education Sciences, Department of Education

The purpose of the initiative is to explore how the federal government can most effectively use its resources to advance the development and effective use of rigorous evidence on what works in crime and substance-abuse policy. The initiative is funded independently by the Robert Wood Johnson and Jerry Lee Foundations. While this final report reflects extensive input from and discussion with the participating federal officials and their staffs, its final conclusions and recommendations are those of the Coalition for Evidence-Based Policy.

II. Proposed strategy: This report proposes a major federal strategy to --

- **Build the knowledge base of crime/substance-abuse interventions shown effective in randomized trials** – not just in demonstration projects but when replicated in community settings; and
- **Spur the widespread use of such evidence-backed interventions by recipients of federal crime/substance-abuse funding.**

The remainder of this report describes our rationale for this federal strategy as well as the specific recommendations comprising the strategy. The recommendations are all designed to be implemented in the near term, and, if coordinated among the federal agencies, can be implemented within the agencies' existing statutory authority and funding levels.

III. The Problem: Crime and substance abuse inflict profound and continuing costs on life and health in America.

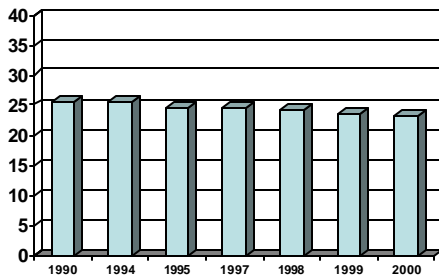
A. Substance abuse – defined broadly to include excess alcohol consumption, drug abuse, and smoking – accounts for approximately 25 percent of U.S. deaths each year.

Specifically, roughly 110,000 U.S. deaths each year are attributable to excess alcohol consumption,¹ at least 16,000 are attributable to illicit drug use,² and approximately 430,000 are attributable to smoking.³ In addition, Americans are victims of over 5 million violent crimes each year, including 15,000 homicides.⁴

B. The United States has made no significant progress in decreasing substance abuse over the past decade. Specifically, the progress our country made in the 1980s against substance abuse ended around 1990, and we have made no significant advances since then. The following charts show the trends for the general population (left column) and youth (right column).

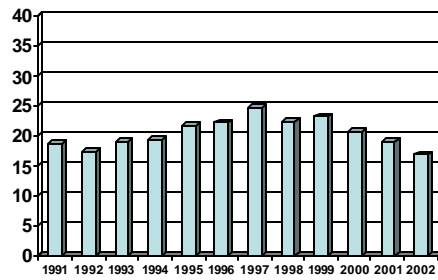
Smoking

Percent of Persons Age 18+ Who Are Current Smokers



Source: National Center for Health Statistics, *Health, United States, 2002*

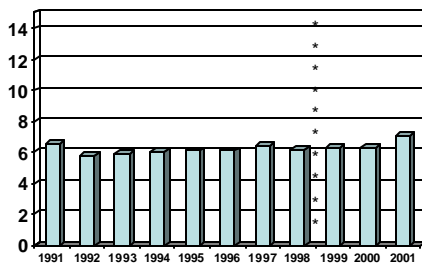
Percent of 12th Graders Who Smoke Cigarettes Daily



Source: National Institute on Drug Abuse, *Monitoring the Future, 2002*

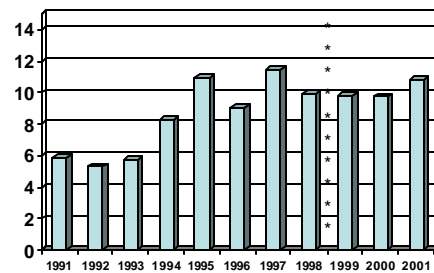
Illicit Drug Use

Percent of Persons Age 12+ Using Illicit Drugs in Past Month



Source: Substance Abuse and Mental Health Services Administration, *National Household Survey on Drug Abuse, 2001 and 2002* (*Survey methodology changed from paper to computer starting 1999, so 1999 and subsequent data are not strictly comparable with prior years.)

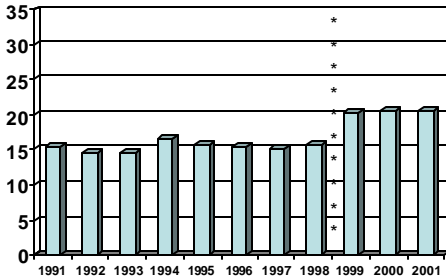
Percent of Adolescents Age 12-17 Using Illicit Drugs in Past Month



Source: Substance Abuse and Mental Health Services Administration, *National Household Survey on Drug Abuse, 2001 and 2002* (*Survey methodology changed from paper to computer starting 1999, so 1999 and subsequent data are not strictly comparable with prior years.)

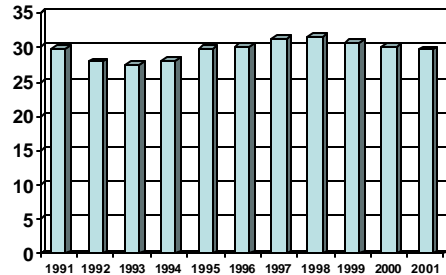
Excess Alcohol Consumption

Percent of Persons Age 12+ “Binge” Drinking in Past Month



Source: Substance Abuse and Mental Health Services Administration, National Household Survey on Drug Abuse, 2001 and 2002 (**Survey methodology changed from paper to computer starting 1999, so 1999 and subsequent data are not strictly comparable with prior years.)

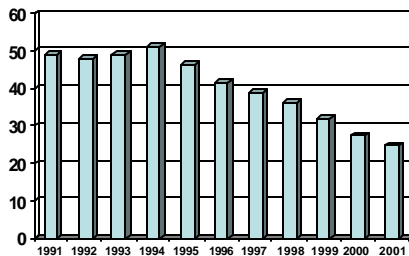
Percent of 12th Graders “Binge” Drinking in Past 2 Weeks



Source: National Institute on Drug Abuse, Monitoring the Future, 2002

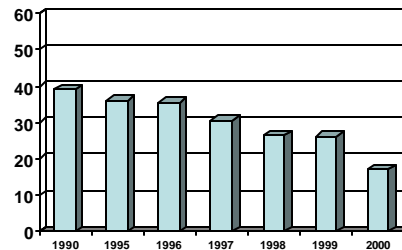
C. By contrast, the United States has made significant progress in reducing violent crime over the past decade, as the following charts show. However, evidence discussed in the next section suggests the crime policy based on rigorous evidence could yield even greater reductions in the future.

Violent Crime Victimization Per 1000 Persons Age 12+



Source: U.S. Department of Justice, Bureau of Justice Statistics, National Crime Victimization Survey

Violent Crime Offenses By Youth Age 12-17 (offenses per 1000 youth)



Source: Federal Interagency Forum on Child and Family Statistics, America's Children: Key National Indicators of Well-Being, 2002, using data from Bureau of Justice Statistics' National Crime Victimization Survey and Federal Bureau of Investigation's Uniform Crime Reporting Program, Supplementary Homicide Reports.

IV. Progress is often thwarted by government programs and strategies (“interventions”) that are not based on rigorous evidence, and research that is not scientifically rigorous.

A. Some of the most widely-used crime/substance-abuse interventions have been shown to be ineffective or harmful in randomized trials – research’s “gold standard” for assessing what works. For example:

- The nation's most widely-used school-based drug use prevention program -- Drug Abuse Resistance Education (DARE), which receives \$200 million annually in public support and operates in 75% of U.S. school districts -- has been shown in randomized controlled trials to have little or no effect on drug use by program participants.⁵ Finally, after 17 years of

operation, DARE announced in early 2001 that it would change its curriculum in response to these studies.⁶

- The nation's juvenile justice system frequently places severely delinquent adolescents in group homes or other congregate care settings – a practice that actually appears to backfire. Randomized trials have shown that such group treatments may *increase* adolescent problem behavior and negative life outcomes, possibly because in a group setting deviant behavior receives positive reinforcement from peers.⁷
- Another crime prevention program that appears to backfire is Scared Straight, in which at-risk or delinquent children are brought into prison to participate in a realistic and confrontational rap session run by prisoners serving life sentences. A recent review of nine randomized controlled trials of Scared Straight and related programs found that these programs either did not affect, or in some cases actually caused a small *increase* in, subsequent criminal activity by program participants.⁸

B. The vast majority of existing crime/substance-abuse interventions, however, have never been rigorously evaluated, and no one knows how effective they are.

- That is a central conclusion of DOJ's 1997 report to Congress, authored by Dr. Lawrence Sherman and others, which reviewed over 500 impact evaluations in the area of crime prevention. To quote the report – "The effectiveness of most crime prevention strategies will remain unknown until the nation invests more in evaluating them. That is the central conclusion of this report. The inadequacy of that investment to date prevents a judgment for or against the effectiveness of the \$3 billion in federal crime funds, at least to a reasonable degree of scientific certainty . . . By scientific standards, there are very few 'programs of proven effectiveness.'"⁹
- Similarly, in the area of youth violence and substance-abuse prevention, a systematic review of over 600 interventions by the respected Blueprints Initiative at the Center for the Study and Prevention of Violence at the University of Colorado at Boulder, has identified only 11 that have been found effective in randomized trials or well-matched comparison group studies.¹⁰
- In the area of drug control enforcement, where the federal government spends \$12 billion annually, a National Academy of Sciences report in 2001 found that there exists a "woeful lack of investment in programs of data collection and empirical research that would enable evaluation of the nation's investment in drug law enforcement. . . . [B]ecause of a lack of investment in data and research, the nation is in no better position to evaluate the effectiveness of enforcement than it was 20 years ago" ¹¹

C. Randomized trials (the "gold standard") are relatively rare in crime and substance-abuse policy.

Randomized trials are considered the "gold standard" in medicine, welfare, and other fields for determining whether an intervention is effective.¹² This is because the process of randomly assigning a large number of individuals into either an intervention group or a control group ensures, to a high degree of confidence, that there are no systematic differences between the groups in any characteristics (observed and unobserved) except one – namely, the intervention group participates in the intervention, and the control group does not. Therefore, assuming the trial is properly carried out (e.g., with a large enough sample), the resulting difference in

outcomes between the intervention and control groups can confidently be attributed to the intervention and not to other factors.

Unfortunately, randomized trials are relatively rare in crime and substance-abuse policy, as in most areas of social policy. For example, Robert Boruch's review of 305 evaluation studies funded by DOJ's Office of Juvenile Justice and Delinquency Prevention in 1998, only 1 percent were randomized controlled trials. Of 179 evaluation studies funded by DOJ's National Institute of Justice in 1998, none were randomized trials.¹³ Anthony Petrosino's comprehensive search of the literature for randomized trials of crime reduction interventions found only about 10 randomized trials reported each year over the period 1981-1993.¹⁴ By contrast, in medicine approximately 10,000 clinical research articles are published annually based on randomized controlled trials¹⁵ – trials which have produced remarkable improvements in American life and health, as discussed below.

D. There are probably many reasons why randomized trials are rare in crime/substance-abuse policy, but the absence of strong federal support for such trials appears to be a major factor.

In some cases, randomized trials are not carried in this policy area because of resistance from communities, judges, police officials, prison officials, or schools to the idea of randomly assigning people – such as defendants, prisoners, or students – into intervention and control groups. Such resistance is sometimes based on ethical concerns about denying access to an intervention thought to be effective. In other cases, the administrative effort and cost of a well-designed randomized trial can be a barrier to carrying out the trial (as discussed in greater detail in recommendation 2).

But to a large extent, the reason why randomized trials have been far more common in medicine and welfare than in crime and substance-abuse policy may be the difference in federal support for such trials in the respective policy areas. For example, based on a 1962 law, the Food and Drug Administration (FDA) in the early 1960s began requiring peer-reviewed randomized trials demonstrating a pharmaceutical drug's effectiveness before the FDA would allow the drug to be marketed.¹⁶ That policy change, along with parallel support for randomized trials by the National Institutes of Health (NIH), transformed the randomized trial in medicine from a rare and controversial phenomenon – which had first appeared in the medical literature only 15 years earlier (1948) – into the widely-used final standard for assessing the effectiveness of all new drugs and medical devices. Between 1966 and 1995, the number of clinical research articles based on randomized trials surged from about 100 to 10,000 annually.¹⁷

Similarly, in welfare policy, the Office of Planning, Research, and Evaluation within HHS's Administration for Children and Families has, over the past 25 years, consistently funded and facilitated randomized trials of welfare-to-work programs and other employment, income supplementation, and related programs for the poor. That support, along with support for such trials from the Office of Management and Budget (OMB) and the White House in the more recent years, has resulted in the implementation of more than 85 randomized trials in this policy area – many of them large-scale, well-designed trials that provide convincing evidence about the effectiveness of particular programs and approaches.

Randomized trials in medicine and welfare – like those in crime and substance-abuse policy—often pose ethical, administrative, and cost challenges. Yet because of strong federal support, these challenges have been overcome in appropriate cases; a large number of well-designed trials have been carried out; and, as discussed below, these trials have resulted in major improvements in the lives of millions of people.

E. Meanwhile, the most commonly-used *nonrandomized* study designs often produce erroneous conclusions and can lead to practices that are ineffective or harmful.

1. “Pre-post” study designs often produce erroneous results. Pre-post studies examine whether participants in an intervention improve or regress during the course of the intervention, and then attribute any such improvement or regression to the intervention. The problem with this type of study is that, without reference to a control group, it cannot answer whether the participants’ improvement or decline would have occurred anyway, even without the intervention. This often leads to erroneous conclusions about the effectiveness of the intervention. Illustrative examples include the following:

- A recent randomized trial of the San Diego Navy Experiment, a spouse abuse treatment program, found the program to be ineffective. However, most abusers in *both* the intervention and the control group did not commit new acts of abuse. Thus, a pre-post study looking only at whether the intervention group improved over time – without reference to a control group – would have erroneously declared the program a success.¹⁸
- Several of the randomized trials of Scared Straight, discussed above, found a reduction in criminal behavior for both the intervention and control groups following the intervention, with larger reductions for the control group. Thus, in these cases too, uncontrolled pre-post studies would have concluded erroneously that the program was effective when in fact it was harmful.

2. The most common comparison-group study designs also lead to erroneous conclusions in many cases. Specifically:

- **In social policy**, a number of careful “design replication” studies have been carried out to examine whether and under what circumstances comparison-group studies can replicate the results of randomized trials.¹⁹ These investigations first compare participants in a particular intervention first with a control group, selected through randomization, in order to estimate the intervention’s impact in a randomized study design. The investigations then compare the same intervention participants with a comparison group selected through methods other than randomization, in order to estimate the intervention’s impact in a comparison-group design. The difference between the two estimates represents the bias produced by the comparison-group design.

These investigations have shown that most comparison-group studies in social policy (employment, training, welfare-to-work, education) produce biased estimates of an intervention’s effects, because of unobservable differences between the intervention and comparison groups. This is true even when statistical techniques are used to adjust for observed differences between the two groups. In many cases, the bias is large enough that it results in erroneous overall conclusions about whether the intervention is effective, ineffective, or harmful.

- **Examples from medicine also show the important limitations of most comparison-group studies**. A recent, well-publicized example is hormone replacement therapy for postmenopausal women. Over the past 30 years, more than two dozen epidemiological studies (a type of comparison-group study) have found hormone therapy to be effective in reducing the women’s risk of heart disease and stroke. But when hormone therapy was finally evaluated in two large-scale randomized trials – medicine’s “gold standard” – it

was actually found to do the opposite – namely, it *increased* the incidence of heart attacks and stroke, as well as breast cancer.²⁰

Many other important examples exist of medical interventions that initially appeared effective in comparison-group studies, but which were subsequently found in large-scale randomized trials to be ineffective or harmful. For instance:

- enriched oxygen environments for premature infants (found to be harmful);²¹
- beta carotene and vitamin A to prevent lung cancer (found harmful);²²
- idoxuridine to treat herpes encephalitis (found harmful);²³
- bone-marrow transplants for women with advanced breast cancer (found ineffective);²⁴
- angiotensin-converting-enzyme inhibitors to prevent cancer (found ineffective);
- dietary salt restriction to reduce hypertension (found only marginally effective).²⁵

Important examples also exist of interventions that initially appeared ineffective or harmful in comparison-group studies, but which were subsequently proven effective in randomized trials. For instance:

- anti-hypertensive therapy to prevent coronary heart disease;
- aspirin to reduce the risk of major coronary events;
- digoxin for patients with heart failure.²⁶

If randomized trials of these medical interventions had never been carried out and the comparison-group results had been relied on instead, the result would have been needless death or serious illness for millions of people.

F. *Well-matched* comparison-group studies can produce valuable knowledge, but their results, too, need to be confirmed in randomized trials.

The design replication studies discussed above generally support the value of comparison-group designs in which the comparison group is *very closely matched* with the intervention group in the characteristics that the intervention is designed to address (e.g., prior criminal behavior and substance use), demographic characteristics, geographic location, time period in which they are studied, and methods used to collect outcome data. Very few comparison-group studies meet this standard.

Such well-matched comparison-group designs seem to yield correct overall conclusions in most cases about whether an intervention is effective, ineffective, or harmful. However, their estimates of the size of the intervention's impact often contain a substantial amount of bias. As an illustrative example, a well-matched comparison-group study might find that a substance-abuse prevention program reduces the incidence of binge drinking by 40 percent – or, alternatively, by 5 percent – when its true effect is 20 percent. The bias is large enough to lead to incorrect overall judgments about the policy or practical significance of the intervention in a nontrivial number of cases.

Based on these findings, we believe that such well-matched studies can play a valuable role in establishing “potential” evidence of an intervention's effectiveness that merits confirmation in randomized trials. In medicine, carefully-designed epidemiological studies have served this function. For example, the Framingham Heart Study – a well-designed, prospective epidemiological study – has been enormously valuable in identifying promising interventions to reduce heart disease and stroke – interventions which were subsequently tested and validated in

randomized trials. But the evidence cautions strongly against using even the most well-matched comparison-group studies as a final arbiter of what is effective and what is not.

V. The Opportunity: Evidence-based reforms could fundamentally increase the effectiveness of federal crime and substance-abuse policy.

A. Randomized trials have identified a few interventions that are highly effective against crime/substance abuse, suggesting that if policy were based on such findings, it could spark rapid progress.

Randomized trials have identified a few social interventions that are highly effective in addressing the problems of crime and substance abuse. Although rare, their very existence suggests that a concerted government effort to build the knowledge base of these evidence-backed interventions, and spur their widespread replication, could fundamentally improve the effectiveness of federal policy in this area. Illustrative examples of these evidence-backed interventions include:

- Nurse-Family Partnership – a nurse visitation program for low-income, mostly unmarried women during their pregnancy and their children’s infancy. At the 15-year follow-up, children born to nurse-visited women had 56 percent fewer arrests; had 81 percent fewer convictions and probation violations; had 63 percent fewer lifetime sexual partners; smoked 40 percent fewer cigarettes per day; and had 56 percent fewer days drinking alcohol in the previous six months compared to the children of women in the control group.²⁷
- Life Skills Training – a substance-abuse prevention program for junior high school students that teaches social and self-management skills, techniques for resisting peer pressure including drug refusal skills, and consequences of drug use. At the five-year follow-up (end of high school), the program reduced smoking by 20% and serious levels of substance abuse by 30-50%, compared to controls.²⁸
- Prison Therapeutic Community – a program that creates a separate community within a prison for inmates with drug problems who are scheduled for release, offers them counseling and instruction in decisionmaking skills and self-discipline for up to one year after their release, and employs staff who are highly-committed role models of recovering substance abusers. At two-years post-release, the program reduces the rate of reincarceration by 35 percent, compared to controls.²⁹
- Moving To Opportunity – a variation on the federal Section 8 housing voucher program – provides low-income families with vouchers to relocate to low-poverty areas. At the 2-4 year follow-ups, the program reduces arrests of the families’ teenagers for violent offenses by 30-50 percent compared to controls; improves measures of child safety for children in participating families (including attacks, threats, and exposure to guns and drugs) by 80 to 100 percent; and reduces the number of these children’s injuries and accidents by 74 percent.³⁰
- Treatment Foster Care – a program that provides severely delinquent teenage boys with foster care in families trained in behavior management, and emphasizes preventing contact with delinquent peers. One year after the program's end, participants' arrest rates and days incarcerated were more than 50 percent lower than those of a control group assigned to a group home -- the typical treatment for severely delinquent adolescents. The average cost of the treatment was 29 to 47 percent *lower* than the cost of group care.³¹

B. The fields of medicine and welfare policy show that the funding and effective use of randomized trials can produce remarkable advances.

In medicine, randomized trials – and randomized trials alone – have conclusively demonstrated the effectiveness of most of the major medical advances over the past 50 years, including: (i) vaccines for polio, measles, and hepatitis B; (ii) interventions for hypertension and high cholesterol, which in turn have helped bring about a decrease in coronary heart disease and stroke by more than 50 percent over the past half-century; and (iii) cancer treatments that have dramatically improved survival rates from leukemia, Hodgkin’s disease, breast cancer, melanoma, bladder cancer, and cervical cancer. Such advances have profoundly improved life and health in America over the past half-century.

Similarly, in welfare policy, randomized trials funded or facilitated by HHS over the past 25 years have succeeded in building a valuable knowledge base of what works in moving people from welfare to work, and greatly strengthened the effectiveness of federal and state welfare policy.³² For example :

- These trials showed that welfare reform programs that emphasized short-term job-search assistance and encouraged participants to find work quickly had larger effects on employment, earnings, and welfare dependence than programs that emphasized basic education. The work-focused programs were also much less costly to operate.
- The trials showed that welfare-to-work programs often reduced net government expenditures.
- The trials identified a few approaches that were particularly successful (e.g., the Portland, Oregon and Riverside County, California welfare-to-work programs).
- These valuable findings were a key to the political consensus behind the 1996 welfare reform legislation and its strong work requirements, according to leading policymakers who worked on the legislation.³³

Recommendations:

We propose that agency officials participating in this initiative launch a major federal strategy to:

- **Build the knowledge base of crime/substance-abuse interventions shown effective in randomized trials** – not just in demonstration projects but when replicated in community settings; and
- **Spur the widespread use of such evidence-backed interventions by recipients of federal crime/substance-abuse funding.**

We recommend that the officials participating in this initiative – who represent the leadership of the federal agencies primarily responsible for federal crime and substance-abuse policy – chair the implementation of this strategy within their agencies and programs. The specific recommendations comprising our proposed strategy are outlined as follows.

Recommendation 1: That the federal agencies develop a concise, uniform, user-friendly set of principles on what constitutes “rigorous evidence” of an intervention’s effectiveness.

A. Rationale: Currently, the absence of a credible, authoritative set of principles may retard both the development and use of evidence-backed interventions.

In contrast to pharmaceutical medicine, where the FDA has adopted a clear, consistent standard on what constitutes rigorous evidence of effectiveness (discussed below), there currently exists no such standard in crime, substance-abuse, education, or other areas of social policy. This may be a primary contributor to the research problem discussed above – that the vast majority of program evaluations in social policy use study designs that often yield erroneous conclusions about whether an intervention is effective, ineffective, or harmful. The absence of such a standard may therefore be a central reason why we have failed to build a sizeable knowledge base of social interventions demonstrated effective and replicable in randomized trials.

The absence of such a standard also may be a major factor inhibiting the widespread *use* of the few interventions that have been shown effective and replicable in randomized trials. Specifically, potential users of rigorous evidence on what works – including federal, state, and local officials, developers and providers of crime and substance-abuse services, and others – now see a vast array of crime and substance abuse interventions each claiming to be effective, and most claiming to be backed by “studies” showing their effectiveness. Most of these potential users of evidence do not have the expertise to judge whether the cited studies are scientifically rigorous or not, and thus to distinguish the interventions that are truly research-proven from the noise.

In medicine, by contrast, the FDA since the early 1960s has had a concise, user-friendly standard for what constitutes rigorous evidence of effectiveness for a pharmaceutical drug or medical device, based on the randomized controlled trial. The FDA standard is now codified at 21 C.F.R. 314.126. Other medical authorities, such as the NIH and the Institute of Medicine within the National Academy of Sciences, employ a similar standard, based on the randomized trial, in many of their activities.³⁴ As noted early, the FDA and NIH policies have transformed the randomized trial from a rare a controversial phenomenon 50 years ago into the widely-used final standard for assessing the effectiveness of medical interventions.

A main reason why this standard has been so influential in medicine is that the FDA uses it as a primary criterion for determining which drugs and medical devices will be awarded a license to be marketed. Although the federal agencies do not have such regulatory power in crime and substance-abuse policy, they do have other policy levers at their disposal which could imbue an analogous standard in social policy with considerable influence, as discussed in the recommendations below.

B. Proposed Next Step: Federal officials launch a process to develop the principles, drawing on the evidence (above) supporting the importance of randomized trials.

We recommend that the federal officials participating in this initiative convene a short (e.g., 6-month) process for developing a set of basic principles on what constitutes “rigorous evidence” of an intervention’s effectiveness. It may also be desirable to invite OMB and the General Accounting Office (GAO) to designate participants in this process, given their central role in executive branch and Congressional policymaking and policy analysis.

The evidence on the limitations of nonrandomized studies, noted above, provides a strong empirical basis for developing such principles, with the randomized trial serving as the final arbiter of effectiveness. We believe that this set of principles, like the FDA standard, should be a concise, user-friendly document in which the participating federal officials concur.

We recommend that this set of principles address:

- **The types of studies needed to establish “strong” evidence of an intervention’s effectiveness** – for example, randomized trials that:
 - ✓ Are well-designed and implemented (with large sample sizes, low attrition, no differential attrition, valid outcome measures, intent-to-treat analysis, and long-term follow-up);
 - ✓ Demonstrate the intervention’s effectiveness in more than one community setting (either in a large, multi-site trial or separate trials);
 - ✓ Find that the magnitude and breadth of the intervention’s effects are of policy and practical significance.
- **The types of studies that may constitute “potential” evidence of an intervention’s effectiveness** – for example, prospective comparison-group studies in which the intervention and comparison groups are very closely matched (as discussed earlier).
- **Important factors that practitioners should consider when implementing an evidence-backed intervention in their community** – for example, the importance (i) of ascertaining that the intervention has been shown effective in community settings similar to theirs, (ii) of adhering closely to the details of the intervention during implementation; and (iii) of collecting outcome data to confirm that the intervention in fact yields results in their community similar to those found in the studies.

The Coalition for Evidence-Based Policy could help facilitate the agencies’ development of the principles, if desired.

C. The resulting set of principles could greatly accelerate the development and widespread use of evidence-backed interventions.

If the above process succeeds in producing a concise set of principles in which the federal agencies concur, it could greatly accelerate both the creation of a knowledge base of evidence-backed crime/substance-abuse interventions and the widespread use of such interventions. Specifically, the principles would provide – for the first time – a clear, authoritative source of federal guidance that could command considerable influence in contexts such as the following:

- It could be referenced by the federal agencies in their research solicitations and program evaluation efforts, where appropriate, as the principles to which the researcher/evaluator should adhere in designing their studies. The agencies could also provide the principles to the reviewers of proposals in these research/evaluation programs, for their use in the proposal selection process.
- When federal agencies develop lists or web sites of proven and promising interventions, they could use the set of principles as a criterion for deciding which interventions are backed by strong or potential evidence of effectiveness (as discussed in recommendation 4, below).
- Future Congressional legislation in the area of crime and substance abuse could, where appropriate, use the set of principles as a standard to which Congressionally-requested program evaluations should adhere. Also, in cases where Congress seeks to require federally-funded activities to be backed by evidence of effectiveness (as it did in the *No Child Left Behind Act of 2001*, including the provisions governing the Safe and Drug-Free Schools program), the legislation could use the set of principles to define what constitutes such evidence of effectiveness.
- When federal crime/substance-abuse grant programs seek to provide incentives for grant recipients to adopt and use interventions backed by rigorous evidence (e.g., per recommendation 5, below), they could use the set of principles to help define what constitutes “rigorous evidence.”
- OMB’s efforts to evaluate the performance of federal programs in crime, substance abuse, and other areas of social policy – e.g., through its new Program Assessment Rating Tool (“PART”) initiative – could use the set of principles to assess: (i) the effectiveness of agencies’ research programs in building a knowledge base of evidence-backed interventions; and (ii) the performance of agencies’ grant programs in funding evidence-backed activities.

Recommendation 2: That the participating agencies – both individually and together – launch a major strategy to build the knowledge base of evidence-backed crime/substance-abuse interventions.

This recommendation outlines our proposed strategy for agency investment in rigorous studies to build this knowledge base. The next recommendation will discuss how the agencies can finance this strategy.

- A. Given the cost of large randomized trials, we believe it is important for agencies to invest *strategically* in such trials and related research,** so that limited agency funds can generate optimal knowledge about “what works.” What follows is an overview of the cost of such trials and other types of studies.

Large, multi-site randomized trials in community settings – which have the potential, by themselves, to yield “strong” evidence of an intervention’s effectiveness, as discussed earlier – may typically cost in range of \$10 to \$50 million. At the other end, small-scale trials involving

the random assignment of 100-200 individuals – which may contribute to “strong” evidence – can cost in the range of \$300,000 to \$700,000.

Well-matched comparison-group studies – which, as discussed earlier, may yield “potential” evidence of effectiveness – typically cost less than comparably-sized randomized trials, but often not much less. This is because the primary expense in a randomized trial usually is not the random-assignment process but rather the cost of tracking and collecting data on the study participants over a period of several years. In most cases, well-matched comparison-group studies must also incur these costs in order to closely match the intervention and comparison groups on baseline characteristics and on type and quality of outcome data over time.

Comparison group studies which merely analyze pre-existing data sources can cost much less, but are unlikely to meet the criteria for closely-matched groups and therefore, as discussed earlier, to yield results that meet even a threshold for “potential” evidence of effectiveness.

B. Given these costs, we recommend the following guidelines for an agency’s (or program’s) strategic investments in randomized trials and related research.

1. The agency should first identify high-priority areas for building the knowledge base of evidence-backed interventions.

The agency could develop a list of high-priority areas based on such factors as:

- The likelihood of identifying highly-effective interventions in a given area that would substantially improve life outcomes for large numbers of people. For example, an agency might designate school-based substance-abuse prevention as a high-priority area if it believes – based on research about the risk factors in substance abuse and/or previous randomized trials and well-matched comparison-group studies – that further investment in rigorous studies is likely to identify highly-effective interventions that will greatly reduce substance abuse among adolescents.
- A large, ongoing public expenditure of funds in the area, the impact of which could be greatly increased if effective interventions are identified. As an illustrative example, DOJ might choose to designate policing strategies as a high-priority area, given the large, ongoing federal, state, and local investment in police services, the impact of which might be greatly increased by identifying effective, evidence-backed policing strategies.

2. The agency should invest initially in small randomized trials and well-matched comparison-group studies, to build the base of interventions backed by “potential” evidence in these high-priority areas.

These initial studies can test the effectiveness of both new and existing interventions. The purpose is to build a pool of interventions, backed by “potential” evidence, that are promising candidates for larger, more costly randomized trials – the kind needed to establish “strong” evidence of effectiveness. Without such initial studies, the agency could waste scarce resources on large trials of interventions that turn out not to be effective, and that do not contribute to our knowledge of “what works.”

Other things equal, if these initial studies are randomized trials rather than well-matched comparison-group studies, they will yield stronger evidence of effectiveness. But, as discussed earlier, in some cases trials may not be practical because of cost, administrative, or

ethical reasons, or because of resistance from local officials or communities. In these cases, well-matched comparison-group studies may be a desirable alternative.

3. The agency should invest in large randomized trials only in areas where previous evidence suggests likely success, and should ensure the trials are well-designed and implemented.

Specifically, when the agency undertakes a large trial, it should do so in areas where previous studies have identified promising candidates, as discussed above. Furthermore, the agency should make sure that the trial is well-designed and implemented, so that it yields strong evidence on which interventions are effective in which communities and under what conditions – findings that are not compromised by problems such as high attrition or absence of data on long-term outcomes.

C. We recommend that the participating agencies also undertake *coordinated* initiatives to invest in randomized trials and related research.

This recommendation is based on the fact that “research” agencies such as NIMH and NIDA have large research budgets but do not have grant programs that fund state and local crime/substance-abuse services. By contrast, “service” agencies, such as DOJ, ED, and SAMHSA, have much smaller research budgets but large grant programs that fund such services. (The various agency research budgets are detailed in recommendation 3, below.)

Thus, in selected areas, the research and service agencies could usefully undertake coordinated initiatives to implement new interventions in community settings and evaluate them in randomized trials (or, if not possible, well-matched comparison-group studies). Such coordinated initiatives might work as follows. A research agency would launch a research program that awards funds to researchers to implement new interventions in state and local crime/substance-abuse programs and to evaluate those interventions in randomized trials. The service agencies would provide incentives for their state and local grantees to participate in these trials – incentives such as additional grant funds or a competitive preference in grant competitions. Such a coordinated initiative would thus provide a strong incentive for the researchers on one hand and the state and local service providers on the other to join forces to test new interventions in randomized trials.

We recommend that these coordinated initiatives follow the same guidelines for strategic investment as those described above for the individual agency initiatives (e.g., invest in large trials only where previous evidence suggests likely success).

Recommendation 3: That each agency focus the following funds/resources, to the maximum extent practicable, on the above strategy to build the knowledge base of evidence-backed interventions.

A. In some cases, agencies can use their existing research and evaluation budgets to directly fund rigorous studies. Examples include:

- **Research budgets of NIMH (\$1.35 billion in FY 03), NIDA (\$0.97 billion in FY 03), and the National Institute on Alcohol Abuse and Alcoholism (NIAAA) (0.42 billion in FY 03).**

These three institutes of NIH have large research budgets that are unencumbered by Congressional earmarks. In addition, research on violence prevention falls directly within the mission of NIMH, and research on substance-abuse prevention and treatment falls directly within the missions of NIDA and NIAAA. In the instances where evidence-backed crime and

substance-abuse interventions exist, these institutes have often been major funders of the randomized trials establishing their effectiveness (including Nurse-Family Partnership, Life-Skills Training, Treatment Foster Care, and Prison Therapeutic Community, discussed earlier).

- **Research budgets of the other agencies/programs participating in this initiative.**

These include the research budgets of DOJ's National Institute of Justice (\$60 million in FY 03), ED's Institute of Education Sciences (\$139 million in FY 03), and HHS's Office of the Secretary (\$20.5 million in FY 03). These budgets are substantially smaller in size than those of the NIH institutes. In addition, the National Institute of Justice's budget, per Congressional direction, funds primarily technology development, and the ED and HHS research budgets are primarily focused on research in areas other than crime and substance abuse.

But in appropriate instances these budgets could fund rigorous studies of interventions that, at least in part, address issues of crime and substance abuse. The HHS research budget, for example, helped fund one of the randomized trials establishing the effectiveness of the Nurse-Family Partnership program.

- **Funds that the agencies can set aside for research, evaluation, and technical assistance from their major grant programs.**

Specifically, the authorizing or appropriations language governing many of large federal crime/substance-abuse grant programs allow the agencies to set aside a small percentage (typically between 1 and 10 percent) or dollar amount of program funds for research, evaluation, and technical assistance. Although these set-aside funds must serve competing priorities (e.g., technical assistance, studies of program implementation), the agencies could use a portion of them to fund randomized trials and well-matched comparison-group studies.

Major federal programs with such set-aside authority include the following. The size of the set-aside, in percentage terms and equivalent dollar amount for FY 03, is shown in parentheses:

- DOJ's Local Law Enforcement Block Grants (3%, or \$12 million); Juvenile Justice and Delinquency Prevention State Formula Grants and Discretionary Grants (10%, or \$16 million); Juvenile Accountability Incentive Block Grants (1%, or \$2 million); Violence Against Women Formula Grants (\$5 million); and Community-Oriented Policing Services Grants (3%, or \$27 million);
- ED's Safe and Drug-Free Schools and Communities National Programs (\$2 million);
- SAMHSA's Substance Abuse Performance Partnership Formula Grants (5%, or \$85 million);
- ONDCP's National Anti-Drug Campaign for Youth (\$2 million).

In addition to these statutorily-authorized set-asides, agencies may be able to implement informal set-asides for research and evaluation, which could then fund randomized trials and well-matched comparison-group studies. For example, since FY 1994 the Justice Department has transferred 1 to 3 percent of funds from certain crime prevention programs to the National

Institute of Justice each year, to carry out research and evaluation. This percentage allocation is carried out with the informal approval of the Congressional Appropriations Committees, and has endured through changes in Administrations and in the parties controlling Congress. It has, in some years, yielded more than \$10 million.

B. In addition, the agencies' grant programs can give grant recipients major incentives to use *their* discretionary funds to carry out rigorous studies.

By doing so, the agencies can often leverage a much larger pool of resources than their own research and evaluation funds to carry out randomized trials and well-matched comparison group studies. For example:

1. The agencies' *competitive* grant programs can give a competitive priority and/or additional funding to applicants that structure their projects as randomized trials.

Major federal competitive grant programs that could use such an approach include the following (FY 03 program funding is shown in parentheses):

- DOJ's Weed and Seed Program Grants (\$59 million), and Community-Oriented Policing Services Grants (\$929 million);
- The Education Department's Safe and Drug-Free Schools and Communities National Programs (\$156 million);
- SAMHSA's Substance Abuse Treatment Programs of Regional and National Significance (\$319 million), and Substance Abuse Prevention Programs of Regional and National Significance (\$198 million); and
- ONDCP's National Anti-Drug Campaign for Youth (\$223 million).

2. Agencies' *formula* grant programs can provide additional funds to grant recipients to structure their evaluations as randomized trials.

Specifically, many formula grant programs allow and/or require grant recipients to use part of their grant funds to evaluate the effectiveness of their activities. In these cases, the federal agencies could offer to co-fund rigorous studies with interested grant recipients (the agency's share would presumably come from their research and evaluation budgets, discussed above). Major formula grant programs that allow and/or require grant recipients to undertake evaluations, and which could therefore apply this approach, include:

- DOJ's Byrne Formula Grants, Juvenile Justice and Delinquency Prevention State Formula Grants, Violence Against Women Formula Grants, Residential Substance Abuse Treatment Formula Grants;
- ED's Safe and Drug-Free Schools and Communities State Grants;
- SAMHSA's Substance Abuse Performance Partnership Formula Grants.

3. Agencies can use their authority to waive certain statutory or regulatory requirements as an incentive for grant recipients to undertake randomized trials.

Federal welfare policy provides an important and successful precedent for this approach. Specifically, in the mid-1980s, HHS and OMB initiated a “demonstration waiver” policy, under which HHS waived some provisions of federal welfare law to allow certain state-level grantees to test new welfare reform approaches, but only if the grantees agreed to evaluate their reforms in randomized controlled trials. These demonstration waivers became more frequent in the early and mid 1990s. Between the mid-1980s and mid-1990s, they directly resulted in more than 20 large-scale randomized controlled trials of welfare reform programs. As discussed earlier, these trials, along with many other trials funded by HHS through voluntary partnerships with state and local agencies, built valuable knowledge about “what works” in moving people from welfare to work, and had a major impact on policy.

Importantly, the HHS/OMB policy of linking waivers to a requirement for randomized trials was not specifically authorized or directed in statute, but rather was developed and implemented administratively, through a creative application of HHS’s statutory authority.

Currently, ED and HHS have waiver authority which they could use in a number of grant programs to get applicants to undertake randomized trials. As an illustrative example, the ED’s Safe and Drug-Free Schools State Grants program now requires grant recipients at the state and local level to implement activities that are “based on scientifically-based research.” Given the current scarcity of evidence-backed interventions in this area, it may make sense for the program to waive this requirement for grantees that structure their activities as randomized trials, in order to build the knowledge base of such interventions. This could provide a strong incentive for school districts or even individual schools to undertake small-scale randomized trials of crime/substance-abuse interventions – trials which, some researchers have suggested, could perhaps be carried out by schools at low cost and with minimal administrative burden.

Recommendation 4: That each agency establish, or contribute to, a “What Works” web site that provides authoritative, user-friendly information to practitioners on evidence-backed interventions.

Such web sites could play a vital certifying function, identifying evidence-backed interventions for practitioners who would not otherwise have the expertise or resources to themselves review and evaluate the evidence. The web sites could thereby greatly accelerate the use and replication of evidence-backed interventions by state and local officials and others who administer crime and substance-abuse programs. An agency may wish to establish its own such web site on evidence-backed interventions within the agency’s subject area; alternatively, an agency may wish to work with other agencies to establish a joint web site.

A. Several federal agencies and programs have already made important advances in this area, developing lists or summaries for practitioners of crime/substance-abuse interventions that are backed by evidence of effectiveness. These federal efforts include:

- The *Blueprints Initiative* funded by DOJ and the Centers for Disease Control (on interventions to reduce adolescent violent crime, aggression, delinquency, and substance abuse);
- The *Model Programs Guide and Database*, sponsored by DOJ’s Office of Juvenile Justice and Delinquency Prevention (on delinquency prevention programs and strategies);

- *Preventing Crime: What Works, What Doesn't, What's Promising*, a 1997 report to Congress by DOJ's National Institute of Justice, prepared by Lawrence Sherman (on crime-prevention strategies);
- The web-based *National Registry of Effective Programs*, sponsored by SAMHSA (on interventions to prevent or reduce substance abuse and other related high-risk behaviors);
- *The Community Guide*, sponsored by the Centers for Disease Control (on community, population, and health care system strategies to address a variety of public health and health promotion topics including substance-abuse prevention);
- *List of Promising and Exemplary Programs* of ED's Safe and Drug-Free Schools Program (on school-based interventions that promote safe, disciplined, and drug-free schools);
- The web-based *What Works Clearinghouse* being launched by ED's Institute of Education Sciences (on educational interventions, some of which may address issues of youth violence and substance-abuse prevention); and
- *Safe and Sound*, a publication of the Collaborative for Academic, Social, and Emotional Learning (CASEL), funded by ED's Institute of Education Sciences and Office of Safe and Drug-Free Schools (on social, emotional, and academic learning programs, including drug education and anti-violence programs).

B. We recommend that the agencies undertake the following next steps to develop What Works web sites, building on their existing efforts:

1. The Justice Department should establish a What Works web site on evidence-backed crime interventions.

A What Works web site or list addressing crime interventions as a whole does not currently exist. The Blueprints Initiative, funded by DOJ and the Centers for Disease Control, provides an excellent foundation for this effort by identifying effective youth crime and substance-abuse prevention programs, but should be broadened to include effective crime interventions generally. For example, randomized trials have been carried out to evaluate policing strategies, spouse abuse treatment programs, pre-trial release of defendants, and prison-based treatment programs; in some cases, these trials have identified effective interventions, such as the Prison Therapeutic Community program, discussed earlier. In addition, the information should be made available, free-of-charge, via a user-friendly web site (the full Blueprints reports are currently available only by purchase in hard copy). Consistent with the Blueprints approach, DOJ's What Works web site should show the cost, as well as the impact on crime outcomes, of the evidence-backed interventions it summarizes.

2. Each agency's What Works site should show which of its listed interventions are supported by "strong" evidence under the uniform federal principles developed per recommendation 1, above.

Currently, the various agency lists of effective interventions use different criteria for assessing evidence of effectiveness (e.g., comparison-group studies versus well-matched comparison-group studies versus randomized trials), and also use different categories of demonstrated effectiveness (e.g., "proven," "exemplary," "model," "effective," and "promising" programs). The result is that particular interventions are sometimes characterized very differently on the various lists. This inconsistency can undermine the

certifying function of the lists, confusing readers as they seek to distinguish the interventions that are truly supported by rigorous evidence from the noise.

To address this issue while preserving the healthy diversity of agency approaches, we recommend that each of the agency What Works sites – in addition to using its own method of categorizing interventions – also indicate which of its listed interventions are supported by “strong” evidence under the uniform federal principles developed per recommendation 1. This would provide a uniform, authoritative, empirically-based metric for readers to use to identify evidence-backed interventions, and thus serve as a valuable supplement to the agency’s own categorizations.

Recommendation 5: That agencies’ crime/substance-abuse grant programs require applicants to provide a concrete strategy for implementation of interventions supported by “strong” evidence.

This requirement could apply to applicants in both formula and competitive grant programs (other than research programs). It would use the criteria for “strong” evidence in the uniform federal principles, developed per recommendation 1. The requirement may not be appropriate for some types of programs, such as programs that fund demonstration projects to test new interventions. And, of course, before this requirement can be applied in a particular program, there must exist one or more interventions backed by “strong” evidence in that program area. The best scenario would be the existence of several such interventions in the program area. However, we believe the requirement should be applied even if only one such intervention exists, because doing so will (i) help to rapidly disseminate that effective intervention, and (ii) provide a powerful incentive for providers of other interventions to carry out rigorous studies to establish the effectiveness of their interventions, as well.

A. Each applicant’s strategy would:

- 1. Identify the interventions backed by “strong” evidence that the applicant plans to implement in its activities, and cite the relevant randomized trials and any other supporting evidence.** We believe it is important that the applicant be allowed to choose which interventions to include in its strategy – even if those interventions are not listed on the What Works web sites – provided the applicant can show that the interventions meet the criteria for “strong” evidence of effectiveness.
- 2. Discuss the applicant’s strategy to foster widespread implementation of these interventions *with fidelity* – that is, with close adherence to the details of the interventions.**

As part of this strategy, the applicant should provide measurable, quantitative goals for widespread implementation of these interventions (e.g., number of local police departments that will participate). The applicant should also discuss how it will ensure fidelity of implementation.

Fidelity is an issue whose importance is often not fully appreciated. Details of implementation can sometimes make a major difference in an intervention’s effects. The Nurse-Family Partnership program provides an illustrative example. As discussed earlier, under this program, well-trained nurses provide home visits to low-income, mostly unmarried women during the women’s pregnancy and their children’s infancy. The program has been shown to be highly effective in three well-designed randomized trials carried out in community settings. Fidelity of implementation appears to be critical to the program’s impact. Specifically, one of the randomized trials demonstrated that when the home visits are

carried out by paraprofessionals rather than nurses – holding all other details the same – the program is only marginally effective. Furthermore, a number of other home visitation programs for low-income families, providing different types services, have been shown in randomized trials to be ineffective.³⁵

3. Discuss how the applicant will evaluate, after grant award, whether it is successfully implementing the interventions with fidelity, and whether they are having the desired effects.

For example, the applicant might propose to collect data on implementation (e.g., number of local police departments participating in a policing intervention, and descriptive data on implementation within each department). The applicant might also propose to carry out an informal matched comparison-group study to confirm whether the intervention is having the expected effect (e.g., for the policing intervention, by identifying a comparison group of roughly-matched police departments in jurisdictions with similar characteristics, and collecting outcome data on crime in the intervention and comparison groups). Such a comparison-group study, while perhaps not fully meeting the threshold for “potential” evidence, may nevertheless provide a rough sense of whether the program is working as expected.

B. In the agency’s competitive grant programs, this strategy would be an important competitive selection factor.

C. In formula grant programs, this strategy would be a factor in the agency’s review of state or local grantees’ program plans.

In some formula grant programs – such as SAMHSA’s Substance Abuse Performance Partnership Formula Grants – the agency is authorized by law to conduct a meaningful review of the grantees’ program plans, and to ensure that the plans include key elements determined by the agency. In these cases, the agency can use its leverage to ensure that grantees include in their plans a concrete strategy for the widespread implementation of interventions backed by “strong” evidence.

In other formula grant programs, the agency has little leverage to influence the content of grantees’ program plans. Even in these cases, however, the agency can request that grantees include a strategy for the implementation of evidence-backed interventions. Although the agency’s review of such a strategy would be *pro forma*, the request itself may itself influence grantees to look for evidence-backed interventions on the What Works web sites or elsewhere, and to think strategically about how to implement such interventions.

D. In high-priority areas, the agency could require an *independent* evaluation of whether the applicant, after grant award, is successfully implementing the interventions with fidelity.

Recommendation 6: That each agency undertake a major effort to educate the policy and grantee communities on the value of these evidence-based reforms, and to provide technical assistance to facilitate their implementation.

Successful implementation of the recommendations in this report will entail major changes in policy and practice at the federal, state, and local level. Policymakers and practitioners at all these levels must be convinced of the value of these reforms if we expect them to fund or participate in randomized trials in significant numbers, or undertake serious efforts to implement evidence-backed interventions. In addition, many researchers in this policy area currently do not have experience or

expertise in carrying out well-designed randomized trials and well-matched comparison-group studies; and federal, state, and local officials often lack the expertise to identify and effectively implement evidence-backed interventions.

Thus, we recommend that the agencies – individually and/or together – undertake a major education and technical assistance effort focused on the following communities:

- Federal, state, and local officials who manage or administer programs providing crime/substance-abuse services;
- Federal and state officials who manage crime/substance-abuse research programs; and
- Researchers who carry out studies in this policy area.

How this education and technical assistance effort can be implemented most effectively is an educated guess, and perhaps itself susceptible to rigorous evaluation. What follows are illustrative examples of ongoing education and technical assistance efforts that might provide a starting point for this effort:

- SAMHSA’s National Registry of Effective Programs not only provides web-based descriptions of substance-abuse prevention and related programs supported by varying levels of evidence, but also provides technical assistance to SAMHSA grantees and others on how to implement such programs effectively. The Registry’s web site (<http://modelprograms.samhsa.gov>) provides information on these technical assistance initiatives.
- DOJ’s Bureau of Justice Administration has engaged a contractor with expertise in program evaluation to provide technical assistance to grantees, at the start of their grant project, in building an evaluation design into the project.
- Under grant awards from ED, the Coalition for Evidence-Based Policy is providing education and technical assistance in evidence-based approaches to K-12 education, including: (i) a major forum last fall with leaders of the education policy community, to discuss rigorous evidence as a key to progress in K-12 education;³⁶ and (ii) a two-day training session scheduled for December 2003 with state and local education officials, to provide assistance in identifying and implementing evidence-backed interventions.

¹ Databases of the National Institute on Alcohol Abuse and Alcoholism, at <http://www.niaaa.nih.gov/databases/qf-text.htm> cited in *Substance Abuse: the Nation’s Number One Health Problem*, a report prepared by the Schneider Institute for Health Policy at Brandeis University for the Robert Wood Johnson Foundation, February 2001, pp. 5–51.

² Donna L. Hoyert, Kenneth D. Kochanek, and Sherry L. Murphy, “Deaths: Final Data for 1997,” *National Vital Statistics Report*, National Center for Health Statistics, vol. 47, no. 19, June 30, 1999.

³ Editorial Note – 1997, *Mortality and Morbidity Weekly*, Centers for Disease Control, vol. 46, no. 20, May 23, 1997, pp. 444-451.

⁴ Federal Bureau of Investigation, Supplementary Homicide Reports, 2000, cited on web site of the Justice Department's Bureau of Justice Statistics, <http://www.ojp.usdoj.gov/bjs/homicide/tables/urbantab.htm>. National Bureau of Justice Statistics on Criminal Victimization, 2001, at <http://www.ojp.usdoj.gov/bjs/cvictgen.htm>

⁵ Donald R. Lynam et. al., "Project DARE: No effects at 10-year follow-up," *Journal of Consulting and Clinical Psychology*, August 1999, pp. 590-593. Also, Dennis P. Rosenbaum and Gordon S. Hanson, "Assessing the Effects of School-Based Drug Education: A Six-Year Multi-level Analysis of Project D.A.R.E.," *Journal of Research in Crime and Delinquency*, vol. 35, no. 4, November 1998, pp. 381-412.

⁶ "Antidrug Program Says It Will Adopt a New Strategy," *New York Times*, February 15, 2001, p. A1.

⁷ Thomas J. Dishion, Joan McCord, and Francois Poulin, "When Interventions Harm: Peer Groups and Problem Behavior," *American Psychologist*, vol. 54, no. 9, September 1999, pp. 755-764. Also, the Buddy System study summarized in James C. Howell, editor, *Guide For Implementing the Comprehensive Strategy For Serious, Violent, and Chronic Juvenile Offenders*, National Council on Crime and Delinquency and Developmental Research and Programs, Inc., May 1995, pp. 91-92.

⁸ Anthony Petrosino, Carolyn Turpin-Petrosino, and James O. Finckenauer, "Well-Meaning Programs Can Have Harmful Effects! Lessons From Experiments of Programs Such as Scared Straight," *Crime and Delinquency*, vol. 46, no. 3, July 2000, pp. 354-379.

⁹ Lawrence W. Sherman et. al., *Preventing Crime: What Works, What Doesn't, What's Promising*, A Report to the United States Congress, prepared for the National Institute of Justice, 1997, chapter 10.

¹⁰ The Blueprints Initiative's web site is <http://www.colorado.edu/cspv/blueprints>. Blueprints is funded in part by the Justice Department and the Centers for Disease Control.

¹¹ National Research Council of the National Academy of Sciences, *Informing America's Policy on Illegal Drugs; What We Don't Know Keeps Hurting Us*, edited by Charles F. Manski, John V. Pepper, and Carol V. Petrie (National Academy Press, 2001), executive summary.

¹² See, for example, the Food and Drug Administration's standard for assessing the effectiveness of pharmaceutical drugs and medical devices, at 21 C.F.R. §314.126. See also, "The Urgent Need to Improve Health Care Quality," Consensus statement of the Institute of Medicine National Roundtable on Health Care Quality, *Journal of the American Medical Association*, vol. 280, no. 11, September 16, 1998, p. 1003; and Gary Burtless, "The Case for Randomized Field Trials in Economic and Policy Research," *Journal of Economic Perspectives*, vol. 9, no. 2, spring 1995, pp. 63-84.

¹³ Robert Boruch, Brooke Snyder, and Dorothy DeMoya, "The Importance of Randomized Field Trials," *Crime and Delinquency*, vol. 46, April 2000, pp. 156-180.

¹⁴ Petrosino's work is discussed in Robert Boruch, Dorothy DeMoya, and Brooke Snyder, "The Importance of Randomized Field Trials in Education and Related Areas," in *Evidence Matters: Randomized Trials in Education Research*, edited by Frederick Mosteller and Robert Boruch (Brookings Institution Press, 2002), pp. 50-79.

¹⁵ Mark R. Chassin, "Is Health Care Ready for Six Sigma Quality?" *The Milbank Quarterly*, vol. 76, no. 4, 1998, p. 574.

¹⁶ The FDA policy is codified at 21 C.F.R. §314.126. Its statutory basis is the 1962 Kefauver-Harris Amendments to the Food, Drug, and Cosmetic Act (see esp. 21 U.S.C. Section 355(d)).

¹⁷ Op. cit., no. 15.

¹⁸ Franklyn W. Dunford, "Determining Program Success: The Importance of Employing Experimental Research Designs," *Crime & Delinquency*, vol. 46, no. 3, July 2000, pp. 425-434.

¹⁹ Howard S. Bloom et. al., “Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?” MDRC Working Paper on Research Methodology, June 2002, at <http://www.mdrc.org/ResearchMethodologyPprs.htm>. James J. Heckman, Hidehiko Ichimura, and Petra E. Todd, “Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,” *Review of Economic Studies*, vol. 64, no. 4, 1997, pp. 605-654. Daniel Friedlander and Philip K. Robins, “Evaluating Program Evaluations: New Evidence on Commonly Used Nonexperimental Methods,” *American Economic Review*, vol. 85, no. 4, September 1995, pp. 923-937; Thomas Fraker and Rebecca Maynard, “The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs,” *Journal of Human Resources*, vol. 22, no. 2, spring 1987, pp. 194-227; Robert J. LaLonde, “Evaluating the Econometric Evaluations of Training Programs With Experimental Data,” *American Economic Review*, vol. 176, no. 4, September 1986, pp. 604-620. Roberto Agodini and Mark Dynarski, “Are Experiments the Only Option? A Look at Dropout Prevention Programs,” Mathematica Policy Research, Inc., August 2001, at <http://www.mathematica-mpr.com/PDFs/redirect.asp?strSite=experonly.pdf>. Elizabeth Ty Wilde and Rob Hollister, “How Close Is Close Enough? Testing Nonexperimental Estimates of Impact against Experimental Estimates of Impact with Education Test Scores as Outcomes,” Institute for Research on Poverty Discussion paper, no. 1242-02, 2002, at <http://www.ssc.wisc.edu/irp/>. This literature is systematically reviewed in Steve Glazerman, Dan M. Levy, and David Myers, “Nonexperimental Replications of Social Experiments: A Systematic Review,” Mathematica Policy Research discussion paper, no. 8813-300, September 2002.

²⁰ Denise Grady, “Scientists Question Hormone Therapies for Menopause Ills,” *New York Times*, April 18, 2002. Gina Kolata and Melody Peterson, “Hormone Replacement Study A Shock to the Medical System,” *New York Times*, July 10, 2002, p. A1. Gina Kolata, “In Public Health, Definitive Data Can Be Elusive,” *New York Times*, April 23, 2002, p. F1.

²¹ W.A. Silverman, *Human Experimentation: A Guided Step Into the Unknown* (Oxford University Press, 1985).

²² Charles Marwick, “Trials Reveal No Benefit, Possible Harm Of Beta Carotene and Vitamin A for Lung Cancer Prevention,” *Journal of the American Medical Association*, vol. 275, no. 6, February 14, 1996, p. 422.

²³ Charles A. Alford, Jr. and Richard J. Whitley, “Treatment of Infections due to Herpesvirus in Humans: A Critical Review of the State of the Art,” *Journal of Infectious Diseases*, vol. 133, supplement, June 1976, pp. A101-A108.

²⁴ Gina Kolata and Kurt Eichenwald, “Health Business Thrives on Unproven Treatment, Leaving Science Behind,” *New York Times*, October 2, 1999.

²⁵ J. Swales, “Population Advice on Salt Restriction: The Social Issues,” *American Journal of Hypertension*, vol. 1, part 1, January 13, 2000, pp. 2-7.

²⁶ The examples not otherwise footnoted are discussed in Stephen MacMahon and Rory Collins, “Reliable Assessment of the Effects of Treatment on Mortality and Major Morbidity, II: Observational Studies,” *The Lancet*, vol. 357, February 10, 2001, pp. 455-462.

²⁷ David L. Olds et. al., “Long-term Effects of Nurse Home Visitation on Children’s Criminal and Antisocial Behavior: 15-Year Follow-up of a Randomized Controlled Trial,” *Journal of the American Medical Association*, vol. 280, no. 14, October 14, 1998, pp. 1238-1244. David L. Olds et. al., “Long-term Effects of Home Visitation on Maternal Life Course and Child Abuse and Neglect: 15-Year Follow-up of a Randomized Trial,” *Journal of the American Medical Association*, vol. 278, no. 8, pp. 637-643. David L. Olds et. al., “Home Visiting By Paraprofessionals and By Nurses: A Randomized, Controlled Trial,” *Pediatrics*, vol. 110, no. 3, September 2002, pp. 486-496. Harriet Kitzman et. al., “Effect of Prenatal and Infancy Home Visitation by Nurses on Pregnancy Outcomes, Childhood Injuries, and Repeated Childbearing,” *Journal of the American Medical Association*, vol. 278, no. 8, August 27, 1997, pp. 644-652.

²⁸ Gilbert J. Botvin et. al., “Long-Term Follow-up Results of a Randomized Drug Abuse Prevention Trial in a White, Middle-class Population,” *Journal of the American Medical Association*, vol. 273, no. 14, April 12, 1995, pp. 1106-1112. Gilbert J. Botvin with Lori Wolfgang Kantor, “Preventing Alcohol and Tobacco Use Through Life Skills Training: Theory, Methods, and Empirical Findings,” *Alcohol Research and Health*, vol. 24, no. 4, 2000, pp. 250-257.

²⁹ Harry K. Wexler et. al., “The Amity Prison TC Evaluation – Reincarceration Outcomes,” *Criminal Justice and Behavior*, vol. 26, no. 2, June 1999, pp. 147-167.

³⁰ Lawrence F. Katz, Jeffrey R. Kling, and Jeffrey B. Liebman, “Moving To Opportunity in Boston: Early Results of a Randomized Mobility Experiment,” *Quarterly Journal of Economics*, May 2001, pp. 606-654. Also, Jens Ludwig, Greg J. Duncan, and Paul Hirschfield, “Urban Poverty and Juvenile Crime: Evidence From a Randomized Housing-Mobility Experiment,” *Quarterly Journal of Economics*, May 2001, pp. 655-679.

³¹ Patricia Chamberlain and Kevin Moore, “Models of Community Treatment for Serious Juvenile Offenders,” in *Social Programs That Work*, edited by Jonathan Crane (Russell Sage Foundation, 1998), pp. 258-276. Philip A. Fisher and Patricia Chamberlain, “Multidimensional Treatment Foster Care: A Program for Intensive Parenting, Family Support, and Skill Building,” *Journal of Emotional and Behavioral Disorders*, vol. 8, no. 3, fall 2000, pp. 155-164.

³² See, for example, Manpower Demonstration Research Corporation, *National Evaluation of Welfare-to-Work Strategies: How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs* (U.S. Department of Health and Human Services and U.S. Department of Education, November 2001).

³³ See, for example, statement of Ron Haskins, who in 1996 was the staff director of the House Ways and Means Subcommittee with jurisdiction over the welfare reform bill, in *Rigorous Evidence: The Key to Progress in Education? Lessons from Medicine, Welfare and Other Fields*, Coalition for Evidence-Based Policy, Council for Excellence in Government, November 18, 2002, pp. 67-69.

³⁴ See, for example, “The Urgent Need to Improve Health Care Quality,” Consensus statement of the Institute of Medicine National Roundtable on Health Care Quality, op. cit., no. 12; and the NIH “consensus statements” on the evidentiary basis for particular medical interventions, at <http://consensus.nih.gov/cons/cons.htm>.

³⁵ For example, see Robert G. St. Pierre et. al., “Improving Family Literacy: Findings From the National Even Start Evaluation,” Abt Associates, September 1996. Karen McCurdy, “Can Home Visitation Enhance Maternal Social Support?” *American Journal of Community Psychology*, vol. 29, no. 1, 2001, pp. 97-112.

³⁶ The Coalition’s web site contains a written transcript and webcast video of the forum proceedings, at <http://www.excelgov.org/displayContent.asp?Keyword=prppcEvidence>.