THE EFFECTIVENESS OF COUNTER-TERRORISM STRATEGIES

A Campbell Systematic Review

Cynthia Lum *

Leslie W. Kennedy †

Alison J. Sherley §

January, 2006

* George Mason University, Administration of Justice Program, 10900 University Blvd., MS 4F4, Manassas, VA 20110, clum@gmu.edu. Corresponding Author.
† Rutgers University, School of Criminal Justice and Center for the Study of Public Security, 123 Washington Street, Room 568, Newark, NJ 07102, kennedy@andromeda.rutgers.edu.
§ Rutgers University, School of Criminal Justice and Center for the Study of Public Security, 123 Washington Street, Room 568, Newark, NJ 07102, asherley@pegasus.rutgers.edu.

CAMPBELL SYSTEMATIC REVIEW SUMMARY

Title and Authors

The Effectiveness of Counter-Terrorism Strategies: A Campbell Systematic Review
Cynthia Lum (corresponding author), Leslie Kennedy and Alison Sherley

Objectives

- To determine the effectiveness of counter-terrorism strategies from the available social scientific research literature using systematic review methods.
- To stimulate debate about the cost-effectiveness of expenditures on counter-terrorism measures.

Findings

- There is almost a complete absence of high quality *scientific* evaluation evidence on counter-terrorism strategies;
- What evidence there is does not indicate consistently positive results – some counter terrorism interventions show no evidence of reducing terrorism and may even *increase* the likelihood of terrorism and terrorism-related harm;
- The available evidence suggests:

    - The use of metal detectors in airports reduces hijackings; however, there may also be a substitution or displacement effect of airport security on other types of terrorism (e.g. assassinations, bombings, hostage taking, death and wounded events).
    - Fortifying embassies and efforts to protect diplomats do not appear effective in reducing terrorist attacks on these targets.
    - Increasing the severity of punishment for hijackers does not appear to have a statistically discernible effect on reducing skyjacking incidents although there is very limited research conducted in this area.
    - UN resolutions, without the implementation of metal detectors, have not been shown to reduce terrorism.
    - Retaliatory attacks (for example, the U.S. attack on Libya in 1986 or attacks by Israel on the PLO) have significantly increased the number of terrorist attacks in the short run, particular against the United States, the United Kingdom, and Israel.
    - The existence of intolerant political parties (to terrorism) and the end of the Cold War could increase terrorism events although the findings in this review were uncertain.
    - In the U.S. alone the non-defence costs of homeland security have increased from $9 billion in 2000 to $32 billion in 2005. In light of the uncertain effectiveness of counter-terrorism measures the cost-effectiveness of this expenditure is open to debate.

Caveats and Qualifications

- The available scientific evidence was drawn from only a handful of studies which use moderately rigorous research designs; this limits the strength of evidence and the conclusions that can be drawn from it.
- It may be that further studies of a higher quality will provide stronger evidence of the effectiveness, ineffectiveness, or harm of particular interventions

Implications for Decision Makers

- The evidence base for policy making, strategic thinking and planning against terrorism is very weak; there is an urgent need to commission research and evaluation on counter-terrorism measures to determine whether these strategies work.
- Scientists need to be included in counter-terrorism policy making, strategic thinking, planning and evaluation; some of the secrecy surrounding counter-terrorism activities is unwarranted.

Implications for Research

- More of the research on terrorism and counter terrorism needs to be empirical and evaluative using scientific principles and different types of methodology. Government agencies can assist in this work by providing funding and access to conduct evaluations.

Contents

List of Tables

List of Figures

1 Review Summary

Since September 11th, there have been massive increases in personal, commercial, and governmental expenditures on anti-terrorism strategies, as well as a proliferation of programs designed to fight terrorism. These increases in spending and program development have focused attention on the most significant and central policy question related to these interventions: Do these programs work? To explore research evidence regarding this question, we conducted a Campbell systematic review on counter-terrorism strategies to determine the scope and strength of evaluation research in this area.

In the course of our review, we discovered that there is an almost complete absence of evaluation research on counter-terrorism strategies. From over 20,000 studies we located on terrorism, we found only seven which contained moderately rigorous evaluations of counter-terrorism programs. We conclude that there is little scientific knowledge about the effectiveness of most counter-terrorism interventions. Further, from the evidence we were able to locate, it appears that some evaluated interventions either didn't work or sometimes *increased* the likelihood of terrorism and terrorism-related harm.

The findings of this review dramatically emphasize the need for government leaders, policy makers, researchers, and funding agencies to include and insist on evaluations of the effectiveness of these programs in their agendas. These agendas would include identifying ways to overcome methodological and data challenges often associated with terrorism research, increasing funding to evaluate existing programs through methodologically rigorous evaluation designs, and paying attention to existing evaluations of programs when implementing them. Further, programs should be assessed to establish if they cause more harm than good or if they create unanticipated consequences.

2 Introduction and Background for the Review

Although actual figures of federal, state, municipal, international, and private expenditures are difficult to calculate, there has clearly been a massive increase in government and private spending on anti-terrorism strategies since September 11th (Congressional Budget Office, 2002; 2005; Guinnessy and Dawson, 2002; Issues in Science and Technology, 2002; Macilwain, 2002; Silke, 2004, generally). In 2002, the Congressional Budget Office estimated that since 1998, there would be almost a doubling of "Appropriations for Combating Terrorism and Protecting Critical Infrastructure", given President Bush's 2002 request, from US$7.2 to $13.6 billion dollars (Congressional Budget Office, 2002). Their most recent report states that these appropriations have further increased to $88.1 billion in 2004 (Congressional Budget Office, 2005).

More remarkably, these estimates represent only a subset of United States defense spending and do not include the billions of dollars expended since September 11th on counter-terrorism measures in other U.S. sectors and around the world. Spending has taken a wide array of forms, including the creation of entire government bureaus (such as the Department of

Homeland Security), improving security at airports and borders, increasing research related to responding to biological or chemical agents, creating and enforcing new laws, building or operationalizing prisons to hold newly defined terrorists, improving medical emergency responses, providing aid to foreign countries and developing municipal disaster plans. The House Budget Committee of the U.S. Congress has recently estimated non-Department of Defense funding for homeland security to have risen from $9 billion in 2000 to $32 billion in 2005.[1]

This massive increase in personal, commercial, and governmental expenditures on anti-terrorism strategies, as well as the proliferation of programs designed to fight terrorism, raise arguably the most significant policy question related to these counter-terrorism interventions: Do these programs work? More specifically, do they reduce the likelihood of, or damage from, terrorism events or discourage individuals from acquiring motivations to carry out political violence?

This Campbell systematic review has discovered that the answer to this question is, disturbingly, "we don't know". Of the thousands of studies we located that focused on terrorism, we found only seven which contained evaluations of anti-terrorism strategies (often examining the same strategies) using at least moderately strong evaluation designs to ensure some internal validity and believability in the findings.

Nor did we find any evidence that the minuscule proportion of government counter-terrorism spending related to research has been directed towards evaluating these programs. Only recently, in February of 2004, the Committee on Governmental Reform of the United States House of Representatives probed the issue in a hearing on "Effective Strategies Against Terrorism" before the Subcommittee on National Security, Emerging Threats and International Relations (U.S. House of Representatives, 2004).[2] During those hearings, Representative Christopher Shays stated:

> *Scientists remind us the plural of "anecdote" is not "data." In the realm of national security, a similar axiom would hold the proliferation of counterterrorism strategies does not necessarily mean we are any safer. Only if those strategies guide us inexorably and measurably toward clearly articulated goals will they secure our liberty and property against the threat of a new and dangerous era. (p. 3)*

Yet, despite this promising start, the hearing diverted away from scientific evaluation of anti-terrorism strategies. Ultimately, like many government assessments, the Government Accountability Office (which was asked to conduct the "evaluation" of counter-terrorism programs) was charged to find out whether national strategies had "fundamental characteristics of a coherent strategic framework; one that clearly states a purpose, assesses risk, sets goals, defines needed resources, assigns responsibilities and integrates implementation" (p. 4). Such an assessment lacked the most fundamental requirement of evaluation – connecting programs with measurable outcomes to determine the success of strategies (e.g., reduction in terrorist acts, fear of terrorism, terrorist recruitment, etc). Nor did it appear that the assessment conducted by the

[1] See http://www.house.gov/budget/ .
[2] See http://www.mipt.org/pdf/Effective-Strategies-Against-Terrorism.pdf .

Government Accountability Office was informed by existing evaluation research on anti-terrorism programs.

At least the House Subcommittee had the right idea. With the decisions to commit funds to anti-terrorism strategies comes the responsibility to know whether these programs are effective. Perhaps even more importantly, as Andrew Silke (2005) has recently pointed out, we not only need to determine if measures are ineffective, but also if policies produce harmful effects (e.g., increasing terrorism). As examples in social policy have shown, programs are often not effective and can sometimes increase the problem or cause further social negatives (see McCord, 2003). Further, evaluations of both effectiveness and harm are necessary for policy that involves phenomena like terrorism that can generate irrational thinking and hasty responses. Responsible public policy for these situations requires a more rational approach, especially since the use of terrorism is clearly not a passing fad and continues to be a major challenge. In particular, the choice of programs on which to spend money should depend on whether we have some evidence that they can be successful, or, at the very least, we should seek to evaluate these programs for evidence of effectiveness.

In the scientific community, this type of research has become associated with the evidence-based policy movement. Evidence-based policy suggests that choices to implement intervention programs, like those which attempt to counter terrorism, are based on what we know is effective or that policies should be subjected to evaluations of effectiveness (Cullen and Gendreau, 2000; Davies et al., 2000; MacKenzie, 2000; Nutley and Davies, 1999; Sherman, 1998; Sherman et al., 2002; Weisburd et al., 2003). Evidence-based policy can provide a moderating effect on knee-jerk responses to crises that have become more influenced by fear and moral panics (see Cohen, 1972) than by reason or facts. In both the medical and social sciences, this movement has led not only to a call for increased scientific evaluations on the effects of intervention or treatment programs (and funding for these evaluations), but also for meta-analyses and systematic reviews which serve to make generalizations from multiple studies of similar programs.

To assess the evidence on the effectiveness of counter-terrorism measures, the Campbell Collaboration[3] (see Boruch et al., 2000; Farrington and Petrosino, 2001), through its Crime and Justice Coordinating Group,[4] approved our protocol to conduct a systematic review of research related to anti-terrorism strategies and to comment on the state of counter-terrorism evaluation research. Specifically, this study systematically reviews all terrorism research and literature, searching for evaluations of anti-terrorism policies using the Campbell review framework[5] and meta-analytic techniques. What follows is our analysis and findings. In particular, we highlight not only the research that we did find, but also what we *didn't* find, and make policy suggestions to government agencies and scientists in terms of what is needed.

We anticipated a number of problems and controversies in conducting such a review. First, we expected that despite the proliferation of counter-terrorism and terrorism prevention programs since September 11th, there would be little evaluation research on these policies, making it difficult to locate quality studies to include in the review. Additionally, terrorism as a phenomenon has not been easily or clearly defined (Crenshaw, 1992; Merari, 1991; Wilkinson,

[3] See www.campbellcollaboration.org .
[4] See http://www.aic.gov.au/campbellcj/ .
[5] See http://www.campbellcollaboration.org/Fraguidelines.html .

1986) and therefore, types of strategies could cover a wide range of phenomena and seek a variety of outcomes. Further, there may be strategies that work but might be legally questionable, violate international law, challenge democratic society, or displace problems elsewhere. Other strategies may seem effective but may be in gross violation of human rights norms, such as the inhumane treatment of prisoners or other corrective policies that may not be acceptable in many societies.

Additionally, terrorism and political violence may involve multiple points at which interventions may be effective and where outcomes might be measured. For example, anti-terrorism strategies may include prevention and alleviation of early risk factors, situational prevention of actual events, or post-event responses. Furthermore, because of the rare nature of terrorism events, it also may be difficult to determine whether these strategies actually worked. Thus, researchers may be interested in detecting secondary effects of prevention tactics, such as the reduction of the fear of terrorism or the adequacy and efficiency of the response after an event occurs. Other strategies may focus on detecting potential terrorism events or other high risk situations.

Despite these controversies and challenges, we sought the most comprehensive review possible not only to learn about what we know "works" in terms of reducing terrorism, but also to learn about the state of terrorism research more generally to guide future evaluation research agendas. To do this, we begin by briefly reporting our findings of a comprehensive, more general review of terrorism research. We then report, in detail, the Campbell Systematic Review of evaluation research on counter-terrorism strategies and describe the methodology used to find and analyze evaluations related to anti-terrorism measures.

3 The General State of Terrorism Research

A general review of terrorism research is a helpful beginning for a more systematic review for a few reasons. Because of the anticipated difficulty in locating evaluations of counter-terrorism strategies, an understanding of the distribution of terrorism research can help to focus on these evaluations and target certain areas of research. Secondly, we expected no one academic discipline to have a monopoly on terrorism research and therefore we needed to familiarize ourselves with a variety of perspectives. Similarly, we anticipated that different types of research methodologies might be used to understand subjects related to terrorism which also necessitated a more general survey to understand the depth of the research.

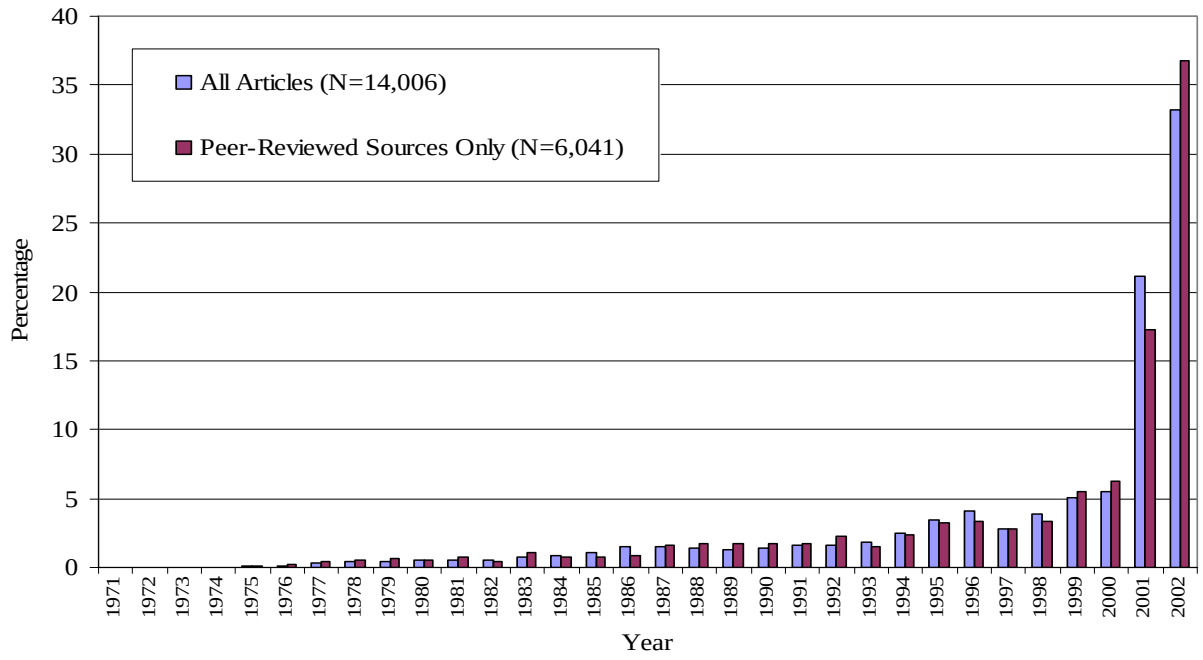
From time to time, a few attempts at reviews of terrorism research have been undertaken (see, e.g. Halkides, 1995; Hoffman, 1992; Miller, 1988; Romano, 1984; Schmid and Jongman, 1988), although September 11th has certainly necessitated the need for an updated review. To conduct this general overview and gain a sense of terrorism research for our systematic review, we collected any article in published, unpublished, peer-reviewed, non peer-reviewed, academic and non-academic sources which mentioned terms related to terrorism and political violence. We conducted this search across seventeen separate literary databases,[6] many of which extend

[6] The databases used were Academic Search Premier, ArticleFirst (OCLC), Contemporary Women's Issues,

back to research conducted since the early 1960s. Although books, government and technical reports, online documents, and web information would be included in our systematic review to be discussed shortly, we limited this initial search to articles in order to gain an overall sense of the literature. We found that there has been much written on terrorism and we located 14,006 articles published between 1971 and 2003 from these databases after excluding duplicate records.

The first, most unique finding compared to past literature reviews was that the events of September 11th have had an enormous effect on terrorism research. As Figure 1 indicates, among the entire 14,006 works located, approximately 54% were published in 2001 and 2002.[7] When only examining articles from peer-reviewed sources (many of the articles were opinion-editorials, news reports, advertisements or general bulletins), this same proportion was found. Neither the Oklahoma City bombing in 1995 nor any other significant terrorist event, including the World Trade Center bombing in 1993, has been followed by this much research interest on terrorism.

Figure 1. Yearly Distribution of Terrorism Publications as a Percentage of Total Publications



Our interest for the Campbell review would ultimately be empirical evaluations of counter-terrorism programs. Thus, to gain a sense of the extent of empirical work on terrorism,

Criminal Justice Abstracts, EbscoHost, EconLit, Educational Abstracts, Electronic Collections Online, ERIC(OCLC), GEOBASE, Humanities Abstracts, Ingenta, ISI Web of Science, MEDLINE, National Criminal Justice Reference Service, PAIS International Articles Only, PUBMEDLINE, Social Science Abstracts, Sociological Abstracts. The time periods covered by each of these databases can be obtained at http://www.lib.neu.edu/gateway/databasestrifold.pdf .

[7] This general review was initiated by the authors in 2003; hence, only literature up to the end of 2002 is represented.

we read each abstract of the peer-reviewed articles to see to what extent empirical analysis had been conducted using a very liberal definition of the term "empirical". Studies deemed empirical had abstracts which indicated that any analysis (either quantitative or qualitative) had been conducted on terrorism data. Thought pieces, on the other hand, were articles where authors discussed an issue theoretically or offered an opinion, while case studies (as deemed by the author) examined a particular situation from a (usually) historical approach.[8]

From this categorization, we found that only 3% of the articles from peer-reviewed sources appeared to be based on some form of empirical analysis. Approximately 1% were categorized as case studies and the rest (96%) were thought pieces. The scarcity of any empirical analysis (whether evaluative or not) on terrorism-related research supported our initial hypothesis that we would find only a small amount of evaluation research on anti-terrorism strategies.

To examine the distribution of specific subject matters studied in terrorism, we grouped each of the articles whose abstract we read into common subject areas. Thirty-five general groups emerged which we collapsed into the seventeen categories shown in Table 1. Table 1 also reports the distribution of these categorizations for studies conducting some form of empirical analysis.

Table 1. Distribution of Subject Matter in Terrorism Research

	Peer- reviewed sources (N=4,458[a])	Empirical only (N=156 [a])
Weapons of mass destruction (biological, chemical, nuclear)	18.1%	10.3%
Article on a specific issue such as the IRA, Al Qaeda or incident [b]	12.2%	5.1%
Political responses to terrorism (war, politics, international relations)	9.5%	1.9%
Causes, motivations, psychology, trends of terrorism	8.7%	18.1%
Impacts of terrorism (political, social, economic)	7.7%	5.2%
Non-political responses to terrorism(medical, social, economic)	5.5%	3.9%
Victimology, coping mechanisms, psychological effects of terrorism	5.4%	25.8%
Other (nationalism, intelligence issues, democracy and vulnerability)	5.4%	3.9%
Legal issues surrounding terrorism	5.2%	0.6%
The media and public attitudes towards terrorism	4.6%	18.7%
How to define terrorism	4.2%	1.3%
Non-conventional, cyber and narco-terrorism	3.0%	0.6%
Religion and terrorism	2.6%	1.3%
State-Sponsored terrorism	2.6%	1.3%
Law enforcement responses to terrorism (airports, police)	2.5%	0.6%
Research/science of studying terrorism	2.1%	0.6%
Domestic terrorism	0.6%	0.6%

[a] Excluding book reviews and articles where not enough information was given to be categorized.
[b] If could not be placed into any other category.

[8] It should be noted that some case studies were based on empirical data, but we categorized them as case studies if specifically indicated by the author.

Generally, issues related to weapons of mass destruction represented the largest proportion of articles (18.9%) followed by articles which focused on a specific issue, such as the Israel-Palestinian conflict, the problems in Northern Ireland, Al Qaeda, or September 11th (if they could not be categorized elsewhere). Additionally, political responses to terrorism and the sociology of terrorism (causes, motivations, explanations, definitions) clearly dominate the research agenda of these studies. These emphases most likely represent the strong political science influence in the study of terrorism.

Surprisingly absent from the literature were articles regarding subjects that are more likely to affect individuals, their fear, the ramifications of anti-terrorism efforts, or the evaluation of policy effectiveness. Generally, there is little written about the effectiveness of law enforcement and other non-political responses to terrorist events. This is also true for issues related to coping with terrorism events, legal concerns, and the general victimology of terrorism. Additionally, when thinking about future threats, the relationships between terrorism and religion, socio-economic factors, and political responses have yet to be made.

When examining those articles preliminarily deemed to be based on the analysis of empirical information, the findings are both encouraging and discouraging. A quarter of the empirical work appeared to be conducted on victimology, a subject seemingly relevant to our search for evaluation studies. While political response literature makes up 9.5% of the literature, only 1.9% of the empirical literature seems to empirically analyze those responses. There appears to be an overrepresentation of empirical literature on the cause and sociology of terrorism, rather than on programs designed to combat it. The vast majority of empirically-based studies did not seem to evaluate the effects of anti-terrorism strategies.

These preliminary findings regarding terrorism research support a number of justifications for a Campbell systematic review. Certainly, as Figure 1 indicates, the study of terrorism is not simply a passing fad of little interest to scientists and evaluation researchers. Given the proliferation of counter-terrorism strategies after September 11th, there is even more reason to increase evaluation research on these programs. Yet, there is a dearth of empirical research on counter-terrorism interventions. This is not to say that the current literature is not useful; the more we know and understand about terrorism and those who wield it, the better. However, much of this literature does not address the effectiveness of counter-terrorism strategies. Nor do we have a grasp on whether measures might be harmful. To explore these concerns, we now turn to the findings from the Campbell review of evaluation research related to anti-terrorism strategies.

4 The Campbell Systematic Review of Counter-Terrorism Strategies

While the general review of terrorism research described above helps to determine the scope and nature of terrorism research, our task was to locate and more specifically examine research evaluating the effectiveness of anti-terrorism strategies. To do this, we narrowed our focus by establishing criteria for the consideration of studies for the review. We then engaged in a systematic search strategy in choosing evaluations deemed at least moderately methodologically rigorous. In the final phase, we extracted information from each study and

used meta-analytic techniques to examine research findings. Each step is now described.

4.1 CRITERIA FOR CONSIDERING STUDIES FOR REVIEW

As the preliminary review of terrorism literature indicated, the objects of study, the research methods used, and perspectives related to terrorism are wide-ranging. The definition of terrorism, and therefore the interventions and measurable outcomes of interventions related to this definition, can be subjective and value-driven. Because of this, a systematic search for evaluations of strategies to counter terrorism can provide a number of challenges not present in other reviews.

Therefore, to provide structure for this review, we centered our search strategy on three organizing constructs. First, we considered what might be deemed terrorism for the purposes of finding evaluations of programs designed to reduce it. This proved challenging because of continuing disagreement over what constitutes terrorism and how it is defined (Crenshaw, 1992; Merari, 1991; Wilkinson, 1986). In terms of definitions, we learned from our general review of terrorism research that what researchers considered as terrorism could range from political violence to domestic violence, rape, or child abuse. As our goal was to initially be as inclusive as possible, we not only considered any study whose concern was politically motivated violence as normally or officially defined,[9] but also any study in which the author(s) referred to the program or outcome as terrorism or terrorism-related, even if not normally or traditionally considered as terrorism or even if their perception of terrorism contradicted official definitions.

The second organizing construct centered on what types of interventions would be considered as counter-terrorism measures. Like interventions intended to reduce criminality or crime, anti-terrorism programs may not directly address the reduction of terrorism events, but rather focus on the reduction of related risk factors. We learned from our preliminary review of the general terrorism literature that what might be considered as an anti-terrorism strategy can include a wide variety of subjects, including political, social, legal, law enforcement, economic, preventative, reactive, or after-care responses. Limiting our search to law enforcement responses, for example, may exclude those responses that focus on psychological coping interventions for victims or situational crime prevention efforts. Additionally, lessons learned from other crime-related systematic reviews (e.g., Sherman et al., 1997; Sherman et al., 2002), show that social intervention programs (e.g., crime prevention programs) are often not carried out by one type of social institution (e.g., the police) but can be carried out by many different institutions (e.g., schools, churches, community groups, etc.). Furthermore, like other types of crime, there may be different time periods prior to an actual terrorist event where interventions might be applied (see Schmid, 1983; Walter, 1969).

Because of the breadth of possible interventions, we initially took the most comprehensive approach possible and considered for our review any study evaluating the effects of programs generally designed to prevent, detect, manage, or respond to terrorism events and related incidents. Prevention strategies can include a wide range of programs designed to deter

[9] For example, the United States Department of State, which has described terrorism since 1983 (Title 22 of the United States Code, Section 2656f[d]) as "premeditated, politically motivated violence perpetrated against noncombatant targets by sub-national groups or clandestine agents, usually intended to influence an audience."

future events, such as improving embassy security, placing metal detectors in airports, or increasing penalties for crimes deemed to be terrorism. Prevention strategies might also include early risk prevention, which could include diplomatic efforts, foreign aid, programs to reduce political or religious fundamentalism, or employment opportunities. Research on detection strategies (which can also be preventative) might include evaluations of measures which improve the scanning of shipping containers, border-related strategies, immigration policy or other tactics used to detect people, places, or situations involved in terror-related activity. Examples of evaluations of management strategies might include examining the government's threat level warning system, new hospital procedures designed to address issues of health security, or programs put in place to help people who have been impacted by a terrorist event, psychologically or otherwise. Evaluations studying response strategies might focus on law enforcement responses to suicide bombing or the effects of economic sanctions on terrorist-harboring states. While many of these categories overlap, we felt that prevention, detection, management, and response strategies provided us with the most inclusive search construct.

The third construct used in considering studies eligible for this review involved deciding what might be viewed as a measurable outcome of anti-terrorism strategies. Because we intended our intervention range to be as comprehensive as possible (prevention, detection, management and response), we broadened the types of acceptable outcomes we would consider from evaluations. The most straightforward outcome measure would be how the intervention affected (hopefully lowered) events related to terrorism. These "events" could be incidents of terrorism, the number of groups who engaged in the use of terrorism, or the frequency of other activities related to terrorist groups. However, other measurable outcomes might include the reduction of the public's fear of terrorism or their feelings of safety from it, the increased ability to respond to certain events, the reduction in the level of risk of an event occurring, or the ability to detect mechanisms of terrorism (for example, the ability to detect anthrax spores in mail). Outcomes measured might include general manifestations (i.e., "terrorism") as well as specific groupings (i.e. "skyjackings", "hostage-taking", "casualty" or "non-casualty").

Outcomes that were too far removed from measurable outcomes with regard to terrorism were not included in the final review. For example, we found one study which used stock market performance as a measure of the effectiveness of assassinations as a counter-terrorism strategy (Zussman and Zussman, 2005a; 2005b). We believe that the use of stock market performance is too indirect a measure of the effectiveness of targeted assassinations of political leaders in reducing terrorism. Further, the study did not offer evidence to support a more direct connection and therefore was not included in the final analysis.

4.2 SEARCH STRATEGY

Using these three general constructs we pursued a systematic and comprehensive search strategy to locate evaluation research. Again, we conducted the most inclusive search possible, given that our preliminary review of terrorism literature suggested a number of challenges in locating research that evaluated anti-terrorism programs. Since we had already read the abstracts for thousands of articles from peer-reviewed sources from our more general review, we examined these studies as a first step in locating evaluations. At this point, we included all studies that made any reference to the evaluation of a program, no matter the methodological

quality of that evaluation. We then went back to the entire database of 14,006 articles we collected and ran multiple keyword searches on their abstracts and titles for words related to evaluations. Any study that included the following key words was extracted:

assess	empirical
assessment	intervention
evaluate	policy
evaluation	program
effect/effectiveness	works

Because the initial collection of 14,006 terrorism studies only included studies up through January 2003 and did not include research from other mediums (books, government and technical reports, online documents, websites, unpublished material), we re-ran our initial search through December 2004 and also extended our search across multiple mediums, including books, government and technical reports, online documents, websites, and unpublished material. Again, the databases that were searched were:[10]

Academic Search	Premier Humanities Abstracts
ArticleFirst (OCLC)	Ingenta
Contemporary Women's Issues	ISI Web of Science
Criminal Justice Abstracts	MEDLINE
EbscoHost	National Criminal Justice Reference Service
EconLit	PAIS (Public/government documents)
Educational Abstracts	PUBMEDLINE
Electronic Collections Online	Social Science Abstracts
ERIC(OCLC)	Sociological Abstracts
GEOBASE	

And, the keywords (and derivations of these words) we used for our initial search included:

anti-Semitic	national security
black supremacist	political crime
bio-terrorism	political violence
bombing	riots
emergency response	skyjacking
ethnic violence	suicide bombing
hijacking	terrorism (all derivatives of the word)
homeland security	weapons of mass destruction
militia group	white supremacist

[10] The time periods covered by each of these databases can be obtained at http://www.lib.neu.edu/gateway/databasestrifold.pdf.

Further, we included combinations of these words with our evaluation keywords listed above to locate evaluation research related to terrorism. We also used labels of significant terrorism events (for example, "September 11th").

Additionally, we searched the internet, as well as terrorism-related organizations and data sources of terrorism research, for studies that might be reported online or unpublished. The complete list of websites and available databases we researched are listed in Appendices A and B. Additionally, when promising articles were located, we searched within articles for possible references to other evaluations. Using this search strategy, we initially located 290 articles, reports, internet publications, and other published and unpublished materials that hinted that some evaluation had been done or was of interest to the study author(s).

We also extended our key word search to multiple book databases using ENDNOTES connection files.[11] We initially discovered 6,415 books that used our terrorism related keywords in their titles or abstracts and after running a key word search on evaluation terms, identified 64 books that might have evaluations within them or refer to evaluations. Combined with our 290 other documents, a total of 354 studies were located that satisfied our initial search for empirically-based evaluation studies. **Thus, of the over 20,000 reports regarding terrorism that we located, only about 1.5% of this massive literature even remotely discussed the idea that an evaluation had been conducted of counter-terrorism strategies.**

There were a number of limitations in this search approach. First, was the authors' dependence on English-language reports. We recognize this is a constraint that cannot be easily overcome in systematic reviews. We did not have contacts within every country to conduct subsequent reviews in multiple languages. Secondly, we did not have access to government reports or data labeled as secret and therefore may have missed classified evaluations (if any) in our search.

4.3 SELECTION PROCESS AND CRITERIA FOR FINAL INCLUSION

The next step was to narrow these 354 works to only those which evaluated anti-terrorism strategies using scientific methods. To do this, we more closely examined the abstracts, titles and notes of each of the 290 non-book studies to see if an evaluation had been conducted, discussed or even proposed or mentioned. In cases where we felt the abstract provided insufficient information, we examined the actual article. For the 64 books, we expected the abstracts to most likely provide insufficient information to help us determine if an evaluation had been conducted or discussed. Thus, we chose to examine each of the 64 books individually at this stage. As with general literature related to intervention programs, it is often difficult to tell from reading the abstracts whether an evaluation had actually been conducted, whether the article was proposing that an evaluation should be conducted, or whether the article simply discussed the process of an intervention (without an evaluation). Thus, using the most liberal approach, we kept any article or book at this stage that even hinted to either conducting or discussing an evaluation of a counter-terrorism strategy.

[11] Using Endnotes version 6, we employed all university library connection files within that version. While this list is too numerous to display here, it can be found at www.endnote.com.

Of the 64 books initially located through our keyword search, 38 books were physically found when searching four separate university libraries and their affiliates.[12] Of these 38 books, only one book indicated that some evaluation of an anti-terrorism program had been undertaken (Wilkinson, 1977). Of the 290 non-book studies initially chosen for examination, 94 abstracts appeared to suggest that the study was at least remotely connected to an evaluation designed to prevent, detect, manage, or respond to terrorism. We were able to physically locate 79 of the 94 non-book studies (84%), again by searching multiple libraries, databases, and the internet, asking for assistance from colleagues and graduate students at different universities, and contacting authors themselves for articles that could not be located. We now had 80 studies that seemed, upon closer examination, to involve some minimal form of evaluation.

The next stage of our selection process involved making a methodological judgment about these 80 studies to determine if they satisfied the minimum requirements to be considered an evaluation. To do this, we chose a two-step process. First, we conducted an initial reading of the full text of each of the eighty studies and chose only those studies which appeared to at least attempt to connect an outcome or effect with a program through a minimally rigorous scientific test. For example, studies which used simple correlation statistics were retained. While this step was a conservative approach, we felt that the nature of terrorism studies (for example, that events are rare or the difficulty in collecting information on terrorism) warranted careful consideration of possible studies to include.

We identified 21 from these 80 studies that satisfied this minimum criterion. Reasons for excluding the other 59 studies were that the study:

- was not an evaluation of a program or policy designed to detect, prevent, respond to, or manage terrorism;
- described the process of a program, but did not evaluate it;
- made claims that a program was effective without any empirical test of that claim;
- was a news article reporting on individuals claiming effectiveness;
- was a review of non-evaluation literature;
- advocated that an evaluation should be done but did not conduct one;
- surveyed individuals about their feelings as to a program's effectiveness;
- surveyed individuals about how prepared they felt for another attack;
- determined the effects of terrorism, rather than the effectiveness of a program;[13]
- examined attrition for participation in a program, but not the effects of the program;
- made suggestions for treatment of terrorism related injuries; and/or
- described criteria as to what effective policy should look like, but did not evaluate anything.

The second step of selecting the final studies for analysis was to only include studies in which authors used more rigorous methodological designs. Not all evaluations are created equal;

[12] Graduate assistants helped search for books at Harvard University, Northeastern University, Rutgers University and the University of Maryland at College Park.

[13] One anonymous reviewer of this review questioned why we did not consider studies examining the effectiveness of terrorism itself. We believed that this literature was generally inappropriate for this review unless the study authors framed their interest in the effectiveness of terrorism in terms of evaluating how terrorism might thwart a counter-terrorism strategy.

they vary in both internal and external validity which lends to the believability of their findings (Cook and Campbell, 1979; Farrington, 2003; Shadish et al., 2002). The minimal requirements of the first stage only indicated that a weak test was conducted connecting the program with the desired outcome. For example, a program designed to treat post-traumatic stress disorder (PTSD) given to victims of terrorist acts may determine that the program worked because scores for indications of PTSD reduced over time after an intervention. However, this finding could be entirely spurious; those individuals who were not given the program may have improved similarly in the same amount of time.

On the other hand, only including randomized controlled experiments in our review would be impractical. Not only would we have nothing to report, but the use of experimentation on rare event outcomes may be too strict a criteria. Thus, as we initially outlined in our protocol, we were guided by what Sherman et al. (1997) describe as a moderate or mid-level scientifically rigorous design. In Sherman et al., they devised a five point "Scientific Methods Scale" (SMS) to score the methodological quality of evaluation research.[14] For them, the highest quality evaluation method, the randomized controlled experiment, was ranked a "5" while simple correlational studies were scored as "1" or "2". We recognize there have been critiques of this scale, and we used it only as a general guide to assist with our selection process.

We used the middle SMS score of "3" as a general guide to exclude those studies which were not at least moderately rigorous, but warn the reader that moderately rigorous designs are less believable than more rigorous ones. Sherman et al. describe an SMS = 3 as "a comparison between two or more units of analysis, one with and one without the program". Throughout Sherman et al. (1997), there are also elaborations on the meaning of a "3" by different authors, suggesting that a methodological score of "3" can also point to studies with multiple units who receive or do not receive the program and attempt to control for other factors.

Time series methods were also coded a score of "3" in Sherman et al.[15] However, as Shadish et al. (2002) point out, time series can be a powerful research design. They equate quality time series designs with quasi-experimental studies. In this study, the vast majority of findings were derived by time series methods and they are viewed as moderate to stronger research designs.

Using this guidance, we decided that our final inclusion of studies from the 21 minimally acceptable studies included those that at least:

- evaluated two or more units of analysis, comparing some with and without the counter-terrorism intervention;
- made some attempt to provide for controls within a statistical analysis;
- conducted an interrupted time series or intervention analysis to indicate some temporal ordering of effects.

After re-reading each of the articles, we determined that 10 studies in Table 2 satisfied our criteria of having used a moderately rigorous methodology in determining the effects of a program on a desired outcome. In Table 2 below, the first column, entitled "Moderately rigorous

[14] The Scientific Methods Scale can be found in Chapter 2 of the Report and can be found at http://www.ncjrs.org/works/chapter2.htm.
[15] This is mentioned in Sherman et al. (1997) on page 2-19, note 7.

method used (Scored at least a "3" on the SMS or equivalent)" denotes these 10 studies.

4.4 FURTHER EXCLUSION OF MEDICAL-RELATED STUDIES

The 10 studies we located that were at least of moderately strong methodological design included three studies from the medical field – Halloran et al. (2002), Lallement et al. (1998), and Quinn et al. (2002). Halloran et al. (2002) created and tested a simulation of the release of smallpox and the effectiveness of targeted and mass vaccinations. They found that smallpox epidemics could be more effectively prevented and controlled with well-timed mass vaccination (prior to smallpox introduction or immediately thereafter) in comparison to targeted vaccination if there is no preexisting immunity. Targeted vaccination, which prevented more cases per dose, was the preferred method if vaccines were limited. Lallement et al. (1998) analyzed the effects of gacyclidine (GK-11) as a therapy for nerve agent poisoning using monkeys in a laboratory setting. They found positive effects of the therapy on reducing the chances of sickness and death associated with nerve agent poisoning. Finally, Quinn et al. (2002) analyzed the effects of an ELISA (enzyme-linked immunosorbent assay) for immunoglobulin G antibodies to anthrax infections. They found that the ELISA was sensitive in diagnosing the presence of anthrax toxins.

Each of these studies used moderate to strong evaluative methods in testing for (or simulating) the effectiveness of these interventions. However, we decided to exclude these three studies for two reasons. First, while our search revealed only three medically-oriented evaluations related to terrorism, we knew that this was not representative of the medical field's research on nerve agents, smallpox or anthrax (or other issues related to terrorism), but rather just what appeared in our search specific to terrorism. For example, a Cochrane review related to Anthrax (Jefferson et al., 1998) already exists. Findings related to these treatments could not be summed up by the three studies we found that happen to focus more specifically on their use in terrorism. As an example, we came across one article during our search – Gillespie et al. (2002) – that didn't make our final cut of having a moderately rigorous design but is a good example, nonetheless. This article discussed the effectiveness of cognitive behavioral therapy on post-traumatic stress disorder (PTSD). Had this article been methodologically rigorous enough to be included, its inclusion in our study would occur only because it directed its study towards victims of terrorism. However, there has been a large literature on the treatment for post-traumatic stress disorder (Bisson and Andrew, 2005; Rose et al., 2002; Stein et al., 2000), and to include one article without including others simply because one referenced terrorism victims would be unfair to the literature.

Secondly, while we generally understood the methodological strategies used by these researchers, the context by which the findings were discussed were much removed from the expertise of the authors. As Lipsey and Wilson (2001) have pointed out, research synthesis often demands a wide range of expertise across fields, and in this particular case, terrorism can extend out towards many different areas of expertise. To interpret the meaning of outcomes measured may require more medical expertise than the authors have. After excluding these three studies, we were left with seven articles in our final selection. Again, we refer to Table 2, which in the last column indicates the final seven studies that were included.

Table 2. Studies Included in the C2 Systematic Review (Last Column) and Final Exclusion Steps

Study listed if it was at least some form of a weak evaluation (SMS=1 or equivalent)	Moderately rigorous method used (Scored at least a "3" on the SMS or equivalent)	Moderately rigorous and not a medical evaluation
Barros, C. (2003). An Intervention Analysis of Terrorism: The Spanish Eta Case. Defence and Peace Economics, 14(6): 401-412.	X	X
Bozzette, S. A., Boer, R., Bhatnagar, V., Brower, J. L., Keeler, E. B., Morton, S. C. et al. (2003). A Model for Smallpox-Vaccination Policy. New England Journal of Medicine, 348(5): 416-425.		
Brophy-Baermann, B., and Conybeare, J. A.C. (1994). Retaliating Against Terrorism: Rational Expectations and the Optimality of Rules Versus Discretion. American Journal of Political Science, 38(1) (Feb): 196-210.	X	X
Cauley, J. and Im, E. (1988). Intervention Policy Analysis of Skyjackings and Other Terrorist Incidents. The American Economic Review, 78(2):27-31.	X	X
Chauncey, R. (1975). Deterrence: Certainty, Deterrence, and Skyjacking. Criminology, 12(4):447- 473.		
Enders, W. and Sandler, T. (1993). The Effectiveness of Antiterrorism Policies: A Vector-Autoregression-Intervention Analysis. The American Political Science Review, 87(4): 829-844	X	X
Enders, W., Sandler, T., and Cauley, J. (1990). UN Conventions, Terrorism, and Retaliation in the Fight Against Terrorism: An Econometric Evaluation. Terrorism and Political Violence, 2(1):83.	X	X
Enders, W. and Sandler, T. (2000). Is Transnational Terrorism Becoming More Threatening? Journal of Conflict Resolution, 44: 307-332.	X	X
Gillespie, K., Duffy, M., Hackmann, A. and Clark, D. (2002). Community Based Cognitive Therapy in the Treatment of Posttraumatic Stress Disorder Following the Omagh Bomb Behaviour Research and Therapy, 40: 345-357.		
Halloran, M. E., Longini Jr., I. M., Nizham, A. and Yang Y. (2002). Containing Bioterrorist Smallpox. Science, 298: 1428-1432.	X	
Johnston, R. G., Garcia, A. R.E., and Pacheco, A. (2002). The Efficacy of Tamper Indicating Devices. Journal of Homeland Security, April. Los Alamos National Laboratory, Vulnerability Assessment Team. Available online at: http://www.homelandsecurity.org		
Lallement, G., Clarencon, D., Masqueliez, C. et al. (1998). Nerve Agent Poisoning in Primates: Antilethal, Anti-epileptic and Neuroprotective Effects of GK-11. Archives in Toxicology, 72: 84-93.	X	
Landes, W.M. (1978). An Economic Study of U.S. Aircraft Hijackings, 1961-1976. Journal of Law and Economics, 21:1-31.	X	X
LeVine, V. T. and Salert, B. A.. (1996). Does a Coercive Official Response Deter Terrorism? The Case of the PLO. Terrorism and Political Violence, 8(1): 22-49.		
Martz, H. and Johnson, M. (1987). Risk Analysis of Terrorist Attack. Risk Analysis, 7(1): 35-47.		
Prunckun, H. and Mohr, P. (1997). Military Deterrence of International Terrorism: An Evaluation of Operation El Dorado Canyon. Studies in Conflict and Terrorism, 20:267-280.		

Quinn, C., Semenova, V., Elie, C., et al. (2002). Specific, Sensitive, and Quantitiative Enzyme-Linked Immunosorbent Assay for Human Immunoglobulin G Antibodies to Anthrax Toxin Protective Antigen. Emerging Infectious Diseases, 8(10):1103-1110.	X	
Smith, B. L., Damphousse, K. R., Jackson, F. and Sellers, A. (2002). The Prosecution and Punishment of International Terrorists in Federal Courts: 1980-1998. Criminology and Public Policy, 1(3): 311-338.		
Smith, B. L. and Orvis, G. P. (1993). America's Response to Terrorism: An Empirical Analysis of Federal Intervention Strategies During the 1980's. Justice Quarterly, 10(4): 661-681.		
Wilkinson, P. (1977). Terrorism and the liberal state. New York: John Wiley and Sons.		
Zussman, A. and Zussman, N. (2005). Assassinations: Evaluating the effectiveness of a counterrorism policy using stock market data. Cornell University, Department of Economics.		

4.5 SUMMARIES OF THE FINAL SEVEN STUDIES

Below are brief summaries for each of the seven reports that satisfied our criteria. As will be detailed, within each of these reports were multiple findings on multiple interventions. However, we summarized the studies here by author(s) as an overview. Additionally, with the exception of Landes (1978) and Barros (2003), the studies primarily use some version of Mickolus's chronologies of terrorism (see Mickolus, 1980; 1982; Mickolus et al., 1989; Mickolus et al., 1989; 1993), which were later used to create the ITERATE databases.

LANDES (1978): Landes' study of United States Federal Aviation Administration data on skyjackings is the earliest evaluation study of counter-terrorism strategies that we could locate and is often referred to in subsequent evaluations. Using analytic techniques common to the field of economics, Landes examined the effects of changes in laws and security measures through the increased probability of apprehension, incarceration, longer sentences, and being killed (by authorities) on the quarterly rate of domestic hijackings and the number of days and flights between successive hijackings for the period between 1961 and 1976. Landes also controlled for other variables, including the number of air flight operations per quarter, civilian unemployment rates, population rates, and personal consumption. He used a variety of ordinary least squares regression techniques to ascertain relationships between these variables. We chose to use Landes' Table 5 (p. 17) as opposed to his Table 3 as it appeared the most complete model in terms of its inclusion of foreign hijackings as a control variable.

CAULEY AND IM (1988): Cauley and Im offered an interrupted time series analysis (also known in economics as intervention analysis) of incidents occurring between 1968 and 1979. In their study, they examine the effectiveness of increased airport security screening measures that occurred in 1973, increased security at embassies and other diplomatic missions in 1976, and the U.N. convention on preventing crimes against diplomatic personnel enacted in 1977. Cauley and Im not only analyzed multiple interventions, but also the effects of these interventions on different outcomes, including skyjackings and non-skyjacking incidents, such as hostage taking, barricades, and attacks on diplomats.

ENDERS, SANDLER AND CAULEY (1990): This evaluation of multiple interventions and

outcomes used an interrupted time series approach for events between 1968 and 1988. Enders et al. examined the effects of metal detectors in airports in 1973, the UN Convention on the Prevention and Punishment of Crimes Against Internationally Protected Persons Including Diplomatic Agents in 1977, United Nations resolutions against General Assembly and Security Council hostage taking (1985), United Nations resolutions against aerial hijacking (1969-1970), as well as the United States retaliatory raid on Libya in 1986.

ENDERS AND SANDLER (1993): Enders and Sandler continued to contribute to the evaluation literature with the use of intervention analysis (and improving upon specific techniques within that analysis) in this 1993 evaluation of the substitution effects that policies may have between different types of terrorism. Like Enders et al. (1990), they examined the effects of metal detectors and resolutions during the period of 1968 - 1988, but also examined security fortification measures taken on U.S. embassies. However, this study differs from their 1990 paper in that they are analyzing interactions and substitution/displacement effects of different interventions across different types of terrorism. Of the many models that Enders and Sandler put forth, we chose to use Model 2 (p. 839) for our review because it had a richer set of variables than Model 1 and did not just include the United States (as did Model 3, p. 841).

BROPHY-BAERMANN AND CONYBEARE (1994): Brophy-Baermann and Conybeare also employ an interrupted time series/intervention analysis approach to determine the effectiveness of six Israeli military-led retaliation attacks on reducing terrorism from the PLO and Lebanon. These retaliations began in September of 1972 in response to the killings of Israeli athletes at the Munich Olympic Games, of which five more retaliations followed through 1988.

ENDERS AND SANDLER (2000): In 2000, Enders and Sandler expanded their analysis to terrorism events between 1970 and 1996. They study the effectiveness of metal detectors, embassy fortification, and the Libyan raid, as well as the reduction in totalitarian governments that occurred after the end of the Cold War. Additionally, rather than study the effects of these interventions on different types of terrorism (or their substitutions) as they did before, they used outcomes which measure the type of person-based destruction. These outcomes included death events, wounded, and non-casualties per quarter. It should also be noted that we considered including a related article in our review (Enders and Sandler, 2002). However, the 2002 article explores alternative time series technical approaches but does not present complete results. The 2002 article is more of an exploration into alternative methods, rather than an evaluation of anti-terrorism interventions.

BARROS (2003): Barros deviates from the use of ITERATE datasets and employs intervention analysis specifically on Spanish terrorist organization, ETA, terrorism using information collected by Abadie and Gardeazabal (2001). The data include assassinations and kidnappings conducted by ETA between 1968 and 2000. He studied the effects of different political ideologies in power, police and military expenditures as well as increases in foreign investment on incidents of kidnappings and assassinations conducted by the Spanish ETA. Barros suggested that the most parsimonious model used a vector autoregression (VAR) time series framework, and thus, we report findings from his Table 6 (p. 410) only. This model only tested for the effects of having a socialist party in power, which he suggests represents “hard line political-

oriented policy" (p. 412), as the increase in expenditures showed no discernible effect in previous models. Unlike the studies above, Barros is the only study which employs a longer time period as the unit of analysis in the time series – years (instead of quarter-years or months).

4.6 Data Extraction, Methodology Used, and Special Issues Regarding Meta-Analyses with Time Series Data

In our Campbell protocol, we proposed to use meta-analytic techniques (see Lipsey and Wilson, 2001) to make generalizations and determine patterns from studies. However, the empirical studies that we uncovered presented unique challenges to the meta-analyses. All of the studies, except for Landes', used intervention analysis with interrupted time series methods (Landes used a related approach). Meta-analyses on time series data is uncommon, and reducing information provided by time series might also over-simplify the findings or knowledge gained.[16] Not only are different time periods analyzed across each of the studies, but, in many cases, different units of analyses are used. For example, 5 of the 7 studies used quarterly time periods (3 month intervals) as their unit of analysis. However, Barros used years as units while Cauley and Im chose months. And, the concept of time (and therefore outcomes associated with it) is qualitatively different than a static unit of analysis.

Thus, although we use many meta-analytic techniques in this review, we emphasize here (and throughout our results) that these techniques are employed for illustrative purposes only and to gain general knowledge about the effects of anti-terrorism strategies. Most importantly, we recognize the limitations of this approach and warn readers that we are making comparisons across different time periods and event categorizations.

As our summaries indicate, each of these studies (with the exception of Brophy-Baermann and Conybeare, 1994) were comprised of multiple findings for different interventions, time periods, and/or outcomes. Many articles reported evaluating multiple interventions within the same study. Some of them also evaluated the effects of multiple interventions on different outcomes, for example, the effect of metal detectors on reducing skyjackings, and the effect of metal detectors on reducing embassy attacks. Additionally, because the studies were interrupted time series, we also had multiple findings across different time periods for the same study. Some authors reported results for both short and immediate time frames as well as long or stable time periods. Thus, in the example of metal detectors, we could have four separate findings within one study. In total, we discovered 86 findings relevant to this review within the seven studies.

One unresolved complication in this meta-analysis is the uncertainty of independent units of analysis. In meta-analysis, analyzed findings should be independent from each other to ensure that assumptions of statistical analyses are not violated when combining findings. For example, short run and long run effects may not be considered independent, but effects of interventions on different types of terrorism might be. In some cases (Cauley and Im, 1988, Enders et al., 1993), substitution or displacement effects are measured from one outcome to another, suggesting non-independence. To gain the most from our data, we chose to code each finding (a strategy suggested by Lipsey and Wilson, 2001), although acknowledge that many report the effects of the same intervention on the same *general* outcome (e.g., "terrorism") but on different specific

[16] David Wilson (George Mason University) in personal correspondence to the authors on June 16, 2005.

outcomes (e.g., "skyjackings", "assassinations", "barricade and hostage situations") and across time periods that may not be independent. Again, we caution the reader that the use of meta-analytic methods on time series research is exploratory and for illustrative purposes only.

To compare these multiple findings, we extracted the effect size from each. When conducting a meta-analysis, the size of the effect must often be standardized across study findings because of the different ways that effects are reported. Although results from time series present aforementioned challenges, one positive aspect of these seven studies was that all of the studies reported their findings in the same way – as the change from the natural rate of events per time period. Thus, because all authors of the studies we selected for review reported their findings as the increase or decrease in terrorism events (however these events were measured), we did not convert the findings to a standardized effect size. In other words, we treated this time series statistic like an unstandardized mean gain of a pre-post contrast (see Lipsey and Wilson, 2001: p. 42), which normally compares the effect of an intervention at two time periods (before and after a treatment is administered).

Many of the findings also provided enough information (t-statistic or standard error of the change) to calculate a 95% confidence interval around the effect size. In many of the cases that reported a long run effect in addition to a short run effect, there was not enough information provided in the long-run effects to calculate a standard error. However, we retained the effect size in our visual displays although we did not include them in our meta-analysis. To calculate an average effect size across study findings, we employed the weighted mean effect size, and its associated standard error and confidence intervals.[17]

We also conducted a homogeneity analysis to test the assumption that there was an underlying population mean that all effect sizes were estimating. Across all findings in which a weighted effect size could be calculated, we rejected the assumption of homogeneity (Homogeneity Q Statistic = 840.25, df = 59). This was expected given the nature of the studies which measured a wide range of outcomes at different time periods. Even within category groupings of specific interventions (see section 5.2), homogeneity was rejected. Thus, when calculating an average effect size for a group of studies, we employed the more conservative random effects model, rather than a fixed effects model.

Given these caveats, for each of the 86 findings, we extracted and/or derived the following information:

- the complete citation
- the specific intervention evaluated (e.g., metal detectors or a military raid)
- the outcome measured as specified by the author (e.g., "skyjacking", "assassination", "deaths"[18])
- the data source the authors used to evaluate interventions
- specific information about the methodology they used
- the unit of analysis that was employed (specifically, the length of the time interval)
- the sample size (i.e., the number of time intervals used in the time series)

[17] See Wilson and Lipsey, 2001: 113-114 for computation formals. Also, see worksheets provided by David Wilson at http://mason.gmu.edu/~dwilsonb/downloads/overview.ppt .

[18] We recognize that incidents such as "skyjackings" and the number of fatalities are two different types of "events". In a skyjacking, there could be multiple deaths. We will specifically note this on visual displays of information as well as on analysis that uses both these types of measures.

- the "effect size" for each finding (i.e., how much the intervention increased or decreased the natural rate of events per time period)
- whether the effect was a short run or long run effect (or whether this was not indicated)
- whether the finding was statistically significant
- the standard error or t-statistic for the change
- the inverse variance weight (to calculate the weighted mean effect size)
- the 95% confidence interval around each effect size.

5 Analysis and Findings

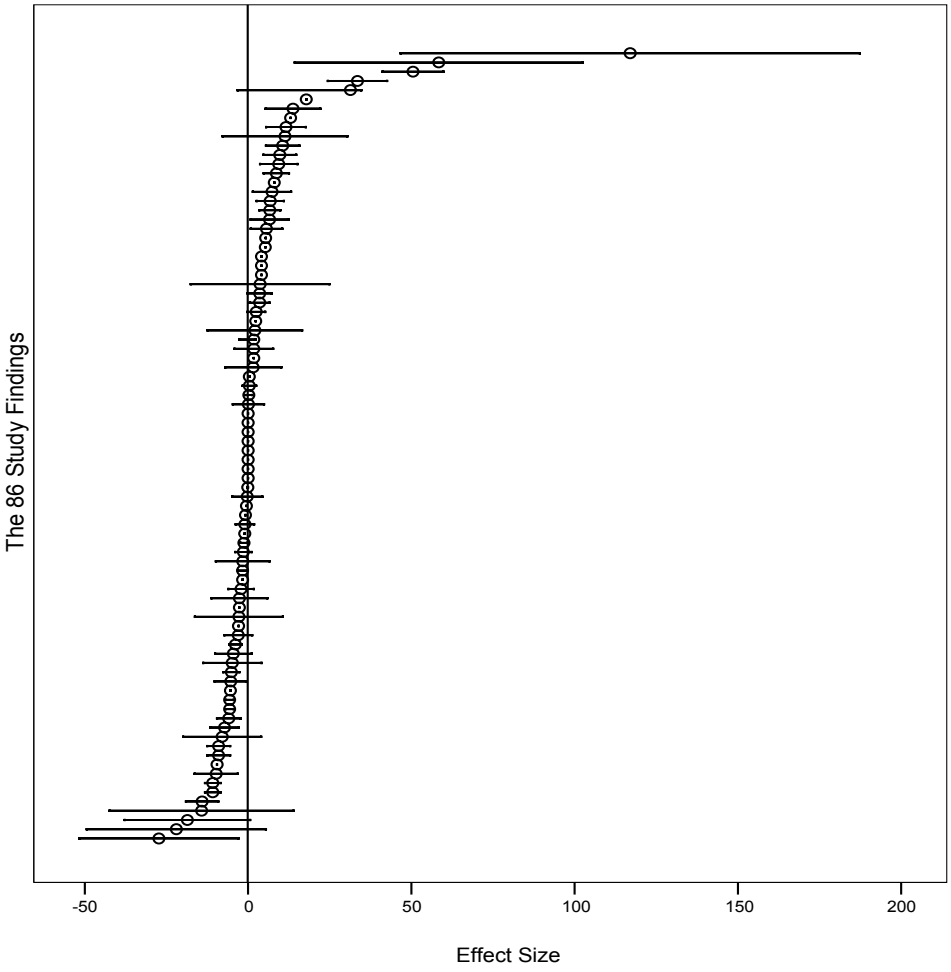
We present our analysis of these multiple findings in two general ways – through visual displays and by analyzing effect sizes across general intervention categories. First, all effect sizes are displayed across all findings, interventions, and outcomes (Figure 2). Secondly, six separate meta-analyses were undertaken for combinations of findings within general categories of interventions, separating out long and short run effects visually (Figures 3-8).

5.1 Overall Effects of Counter-Terrorism Strategies

Figure 2 illustrates the effect sizes and confidence intervals (when calculated) of the 86 findings from the seven studies. Many effects are close to, or cannot be statistically discerned from, a zero effect and therefore provide no statistical evidence that the counter-terrorism strategy was effective or harmful. Additionally, while there appears to be negative effects (which indicate a decline in terrorism events due to the intervention), there are also interventions that appear to have increased the likelihood of a terrorist event occurring (effects greater than zero).

As we warned earlier, summary statistics for all counter-terrorism programs examined can be misleading and difficult to interpret because the values represent effects from different types of interventions, constructs, and time periods. Thus, we cautiously report the following statistics for illustration only: The average of the effect sizes (no matter whether a standard error was reported) was 2.81 events, suggesting generally that across the findings, interventions tended to *increase* the likelihood of terrorism events. However, when using meta-analytic statistics on only those effect sizes which report standard errors, the weighted mean effect size across these findings was 0.30, with a standard error of 0.656, and a 95% confidence interval of {-0.98, 1.59}. For those findings in which standard errors were calculated, there appears to be no statistically significant evidence that interventions had an effect on terrorism. In other words, **across all interventions studied, there does not seem to be a consistent indication of positive effects of counter-terrorism policy.**

Figure 2. Effect Sizes and 95% Confidence Intervals (if available) for All Findings



5.2 EFFECTS OF SPECIFIC INTERVENTION CATEGORIES

The findings in total represent a variety of strategies, outcomes measured and time periods examined. Thus, to better discern the effectiveness of counter terrorism programs, we explored six categories of interventions and their separate effects. For categories with enough findings for a meaningful within-category meta-analysis (more than three findings), we also calculated the weighted mean effect size and confidence intervals. These six categories were:

1. Interventions which increased detection at airports, including installing metal detectors and increasing security screening more generally.

2. Interventions which increased protection, including those which fortified embassies or protected diplomats.

3. Interventions which increased the length and/or severity of punishment for those apprehended and convicted of terrorism.

4. Interventions that represented United Nations Resolutions against terrorism.

5. Military interventions and/or retaliations, specifically, the Israeli retaliation attacks on the PLO and Lebanon in the 1970s and 1980s and the United States attack on Libya in 1986.

6. Changes in political governance, such as having certain political ideologies in power or the end of the cold war (and reduction in totalitarian states).

Figures 3 through 8 illustrate our findings for each of these categories. Each graph is read the same and includes the following attributes:

- The title indicates the general category of intervention in which findings were grouped.

- The citations for each finding are listed on the y-axis. Only 56 (of 86) citations are listed because the long run indicator, if given (see below), is visually displayed directly above its short run counterpart for the same citation.

- The size of the effect is indicated on the x-axis. Effects greater than 0 indicate that terrorism events increased after the intervention and negative effects indicate that terrorism decreased after the intervention.

- A circle with a bar (⊖) represents the effect size and the 95% confidence intervals for the size of the effect when a standard error was available. When the confidence interval crosses 0, this indicates a non-significant effect.

- A stand-alone circle with a dot in the center (⊙) is sometimes given directly above the 95% confidence interval or a cited finding. This stand-alone circle represents the long run finding without a confidence interval in which the author(s) provided both a short run and a long run finding.

- The outcome measured for each finding is indicated next to the effect size and confidence interval.

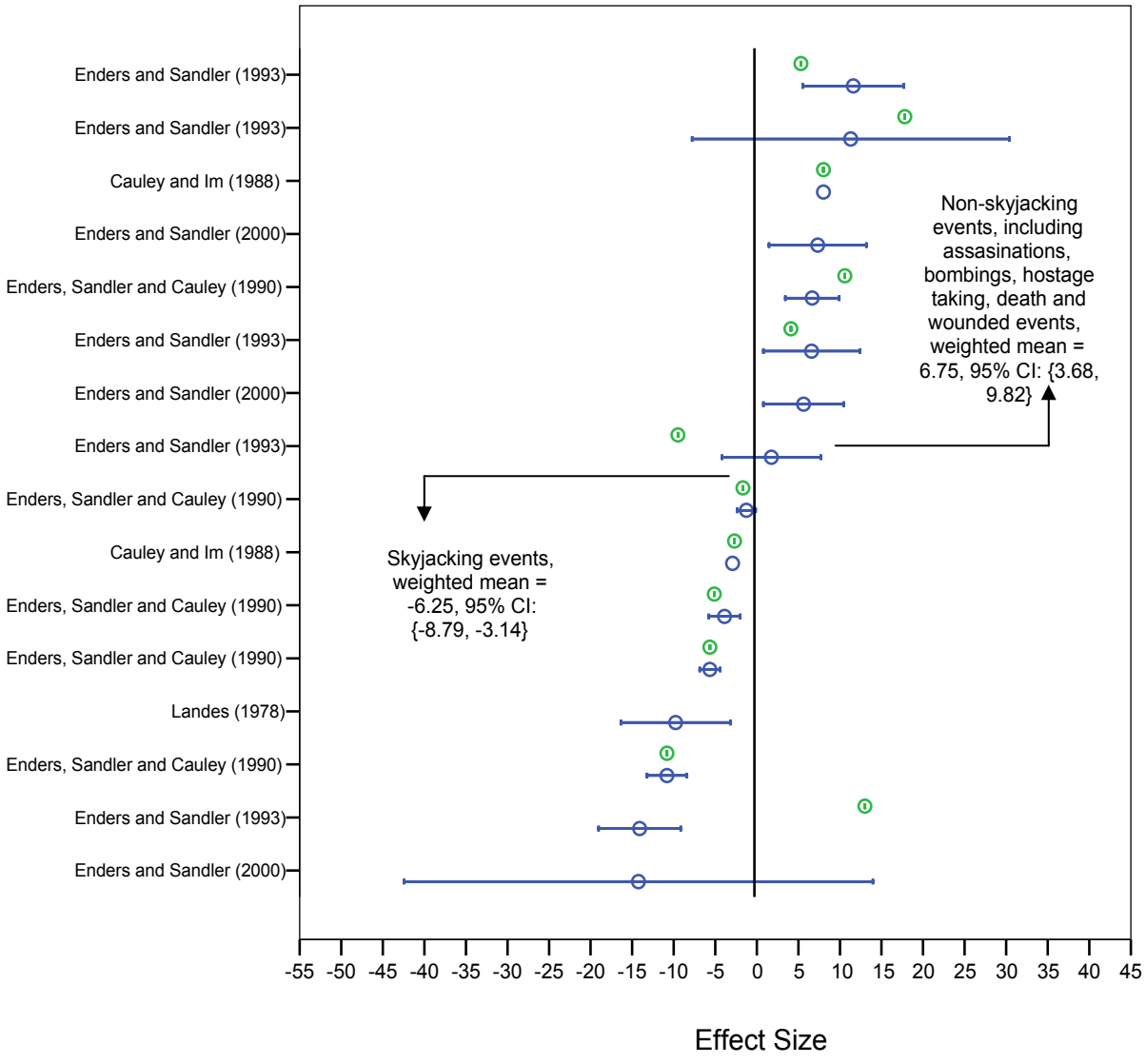
5.2.1 Metal detectors and security screening

We first grouped all findings related to interventions which increased detection of potential terrorism through airport security, one of the most common interventions analyzed by researchers. Specifically, these findings focused on increasing security of airports in the early

1970s including the installation of metal detectors and the more general increase in security screening of passengers.

Figure 3 displays the array of effects related to this intervention. Notice that the general belief of the success of metal detectors is brought into question. First, the success of metal detectors in airports have often been widely recognized as reducing hijackings and the findings here support this for both short and long run measures. The weighted mean effect of airport security on hijacking was a statistically significant reduction of 6.3 events with a 95% confidence interval (CI) of {-8.79, -3.14}.

Figure 3. Increased Detection: Metal Detectors and Security Screening



However, Figure 3 shows an important qualification to the widely held belief of metal detector and airport screening success. For findings where effect sizes were less than zero (the intervention led to a decline in events), all outcomes measured were hijacking events, as indicated in the figure. For those findings indicating a harmful effect (increasing terrorism events), the outcomes measured were non-skyjacking events. As Cauley and Im (1988) and Enders and Sandler (1990; 1993) have repeatedly pointed out, there may have been substitution or displacement effects of airport security on other types of terrorism. In other words, Figure 3 indicates that airport security may decrease airplane hijacking, but can increase miscellaneous bombings, armed attacks, hostage taking, and events which included death or wounded individuals (as opposed to non-casualty incidents) in both the short and long run. The weighted mean effect for these non-hijacking events was a statistically significant *increase* in non skyjacking terrorism by 6.8 events (CI: {3.68, 9.82}).

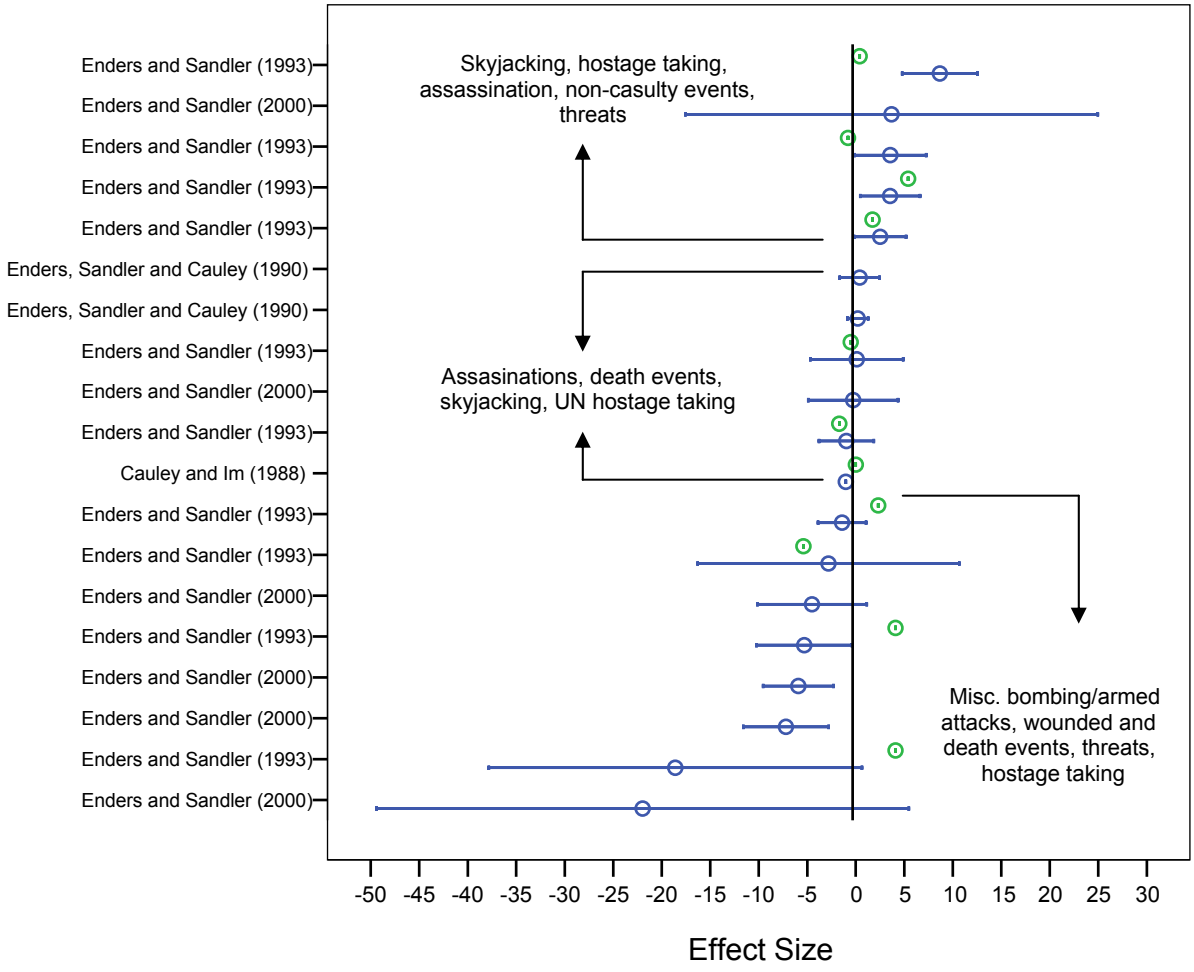
Thus, it appears that while the use of metal detectors reduces hijackings, they may also lead to displacement or substitution effects and inadvertently increase other forms of terrorism. In fact, when calculating the overall weighted mean effect size for all of the findings examining the effectiveness of metal detectors, the positive and harmful effects cancel each other out. The weighted mean (-0.96) is not statistically significant (CI:{-2.91, 0.998}).

5.2.2 Fortifying embassies and protecting diplomats

The second general category by which we grouped findings were evaluations of protective interventions which fortified embassies and protected diplomats, specifically United Nations General Assembly and Security Council members. These findings are reported in Figure 4. As the confidence intervals indicate, many findings were non-significant and close to 0, indicating that there is no scientific evidence these interventions work. Indeed, the weighted mean effect size for these findings was not statistically significant (weighted mean effect size = -0.45, CI: {-2.17, 1.27}). In total, the findings do not indicate that the fortification embassies and efforts to protect diplomats have been effective in reducing terrorist attacks on these targets.

Additionally, there does not appear any logical grouping of the types of outcomes in which harmful or beneficial outcomes occur. When examining individual findings to the right (harmful), left (beneficial) and at the statistically indiscernible area of Figure 4 (around 0 or when a confidence interval of an individual finding crosses 0), no clear pattern emerges. For example, death events can be the outcome of potentially beneficial as well as insignificant counter-terrorism programs. Or, skyjacking events may be the outcome measured in programs that are harmful or have no statistically discernible effect.

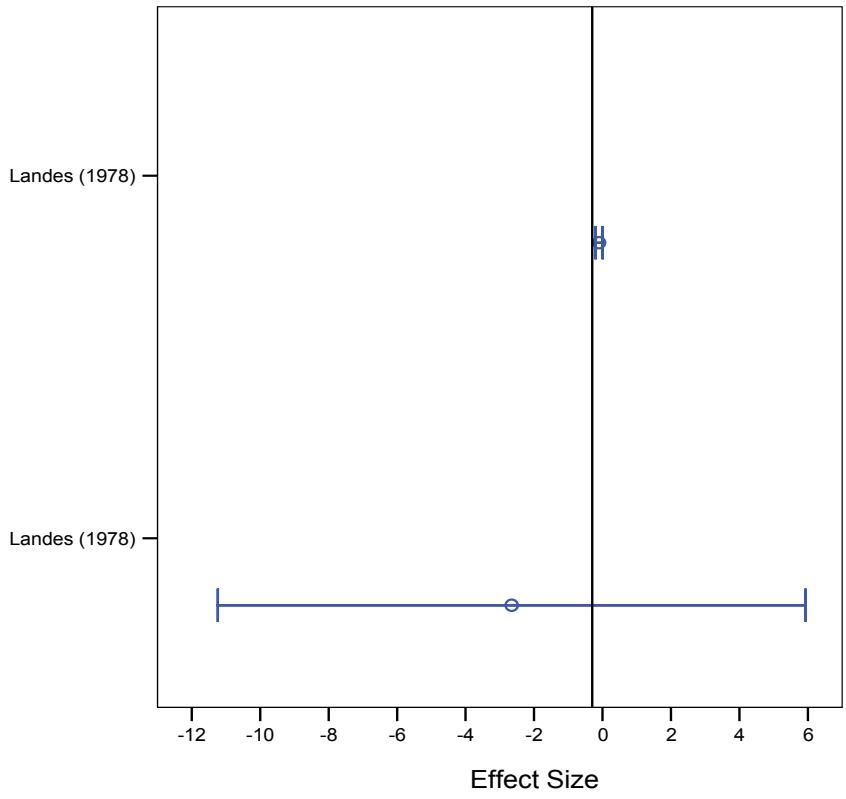
Figure 4. Increased Protection: Fortifying Embassies and Protecting Diplomats



5.2.3 Increasing the Severity of Punishment

Landes (1978) provided the only findings concerning increasing the severity of punishment for hijackers who were apprehended (Figure 5). It does not appear from Landes' work that increasing the severity of punishment had a statistically discernible effect on reducing skyjacking incidents. Again, this does not mean that these strategies "don't work" and methods of time series analysis have become more advanced since Landes' work. However, given the little evidence we have, there appears to be no evidence to show otherwise.

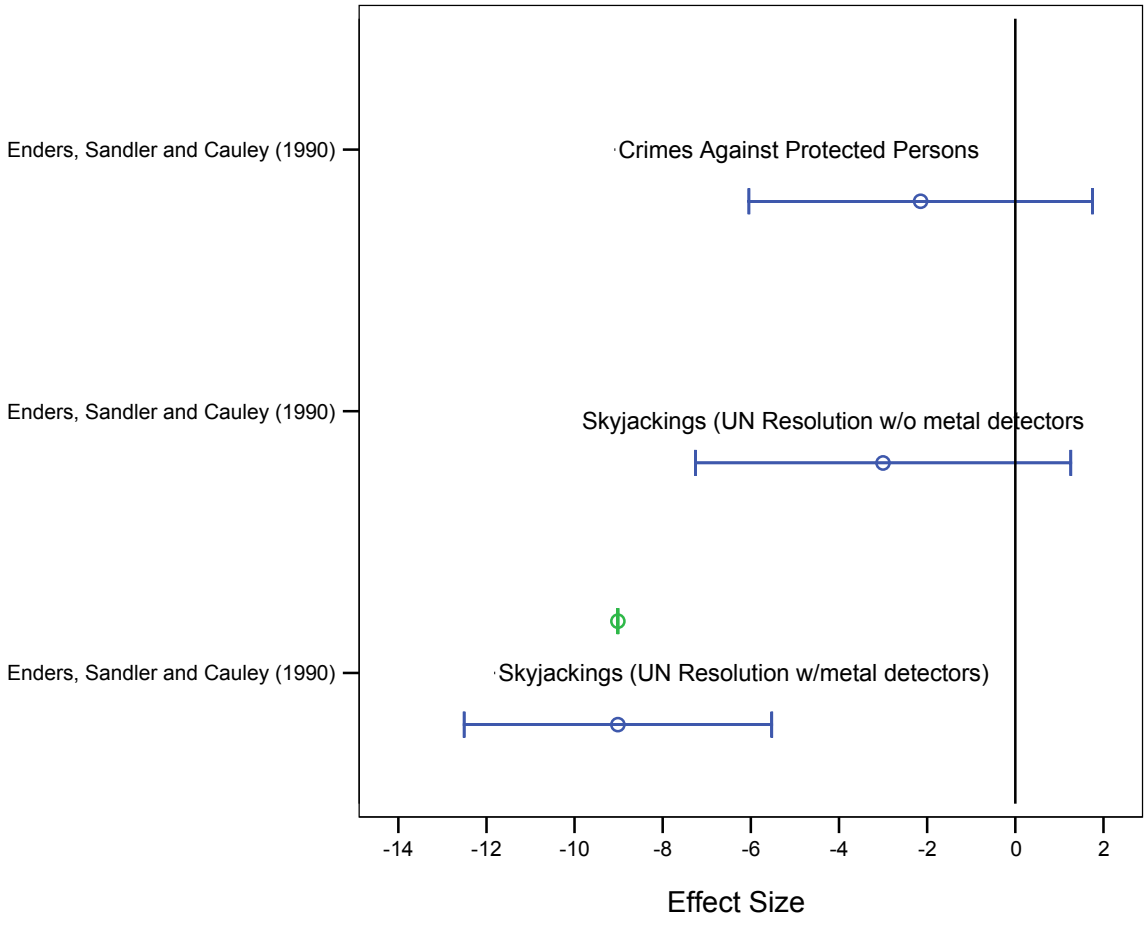
Figure 5. Increased Length of Incarceration and/or Severity of Punishment for Skyjackings



5.2.4 United Nations resolutions against terrorism

Yet another type of intervention found in the evaluation literature concerned the use of United Nations resolutions against terrorism. Although these resolutions are more general in nature, they may provide a general deterrent effect on terrorism by establishing international norms which affect or strengthen national policy against terrorism. As Figure 6 shows, however, only Enders et al. (1990) discovered that only a UN resolution against aerial hijackings that also supported the use of metal detectors in airports appeared effective in reducing the number of skyjacking events in both the short and long term (the finding at the bottom of the graph). However, the second finding illustrated in the middle of Figure 6 indicates that resolutions without the implementation of metal detectors were not useful in reducing terrorism. Further, resolutions intended to "prevent and punish crimes against internationally protected persons" did not seem to have a statistically discernible effect.

Figure 6. Resolution Interventions: United Nations Resolutions Against Terrorism



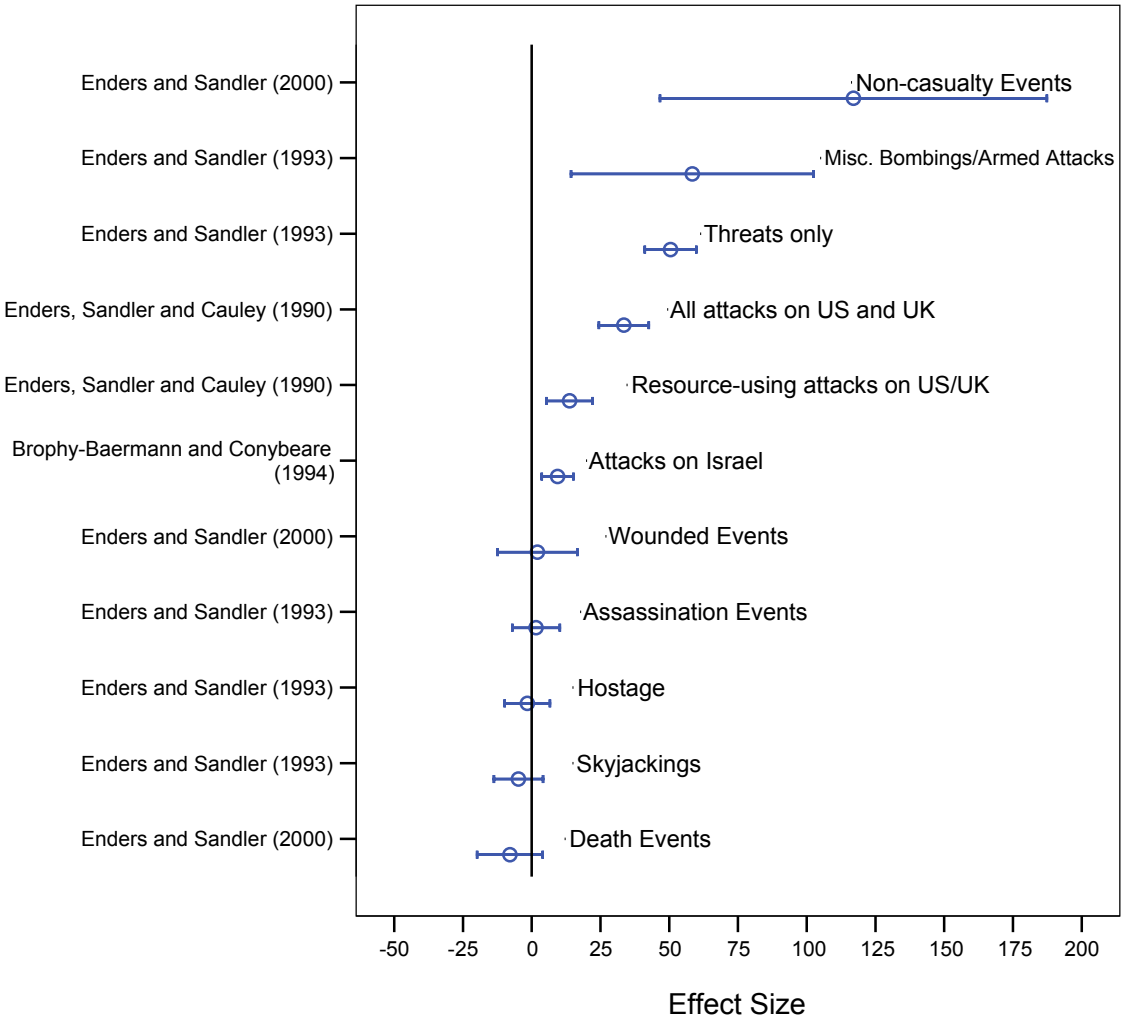
5.2.5 Military Retaliations

We discovered that one often-researched event was the United States' 1986 attack on Libya after Libya's involvement in the bombing of the LaBelle Discotheque in West Berlin. While some have incorrectly reported the effects of the raid as reducing terrorism (see Prunckun and Mohr, 1997), it is generally believed that this raid increased terrorist attacks, at least in the short run (see Silke, 2005).

Again, Enders and his colleagues discovered an interesting nuance in these effects. The findings indicated in Figure 7 point to the short run effects of the attack on Libya (the short run distinction is suggested by the authors themselves, as unlike metal detectors, the attacks on Libya were not sustained over the time series) as well as Israeli retaliatory strikes on Palestinians. Figure 7 generally shows that the attack on Libya resulted in a statistically significant *increase* in the number of terrorist attacks in the short run, with a weighted effect size of 15.33 events, with a 95% confidence interval of {3.46, 27.2}. However, Figure 7 also shows that the Libyan attack affected non-casualty events, threats and miscellaneous bombings more so than "resource-

utilizing" (Enders et al., 1990) attacks such as hijackings, hostage events, and events which lead to death or wounded individuals. Additionally, the specific retaliations researched seemed to increase attacks on the United States, United Kingdom, and Israel.

Figure 7. Military Retaliation by States: Israeli Military-led Retaliation Attacks on the PLO and Lebanon and United States Retaliatory Raids on Libya



5.2.6 Changes in political governance

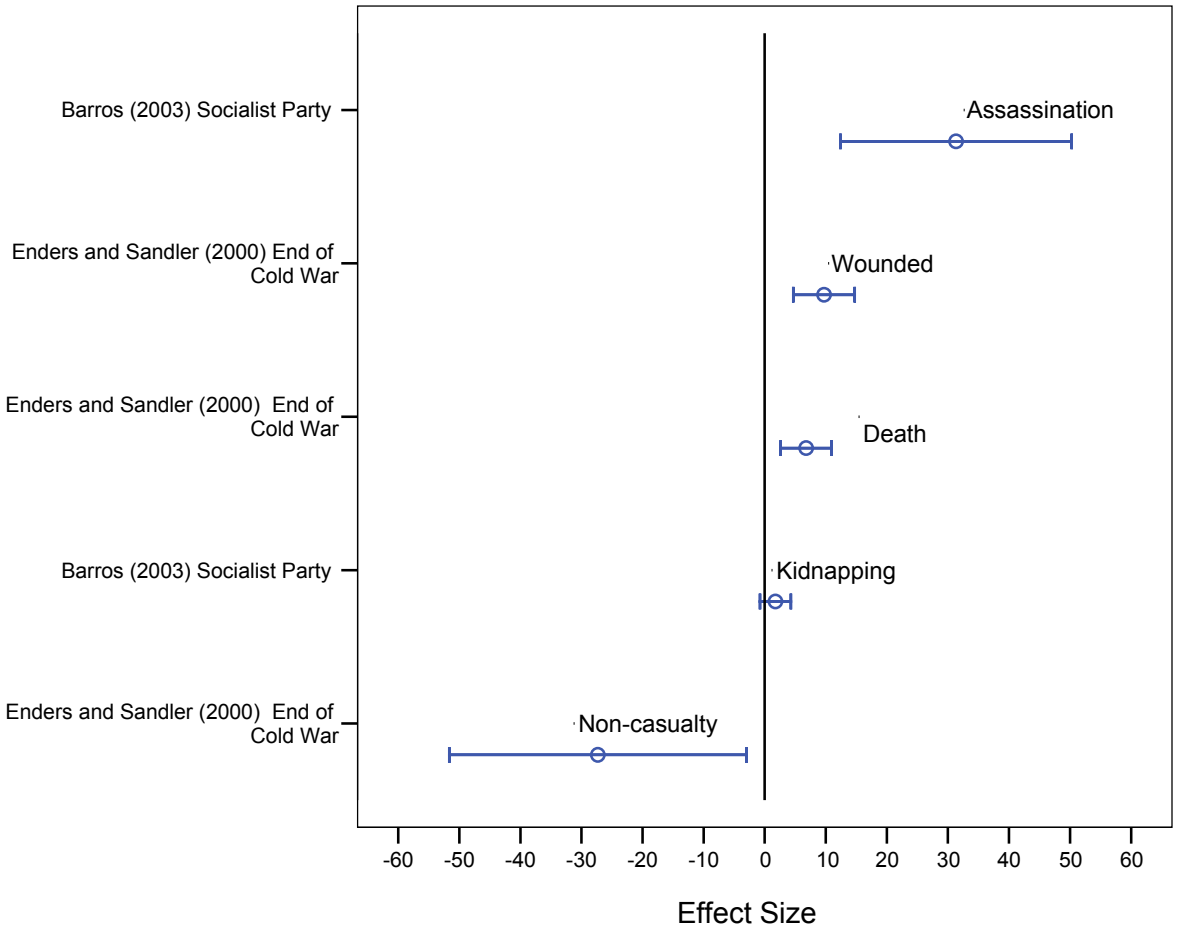
Finally, we examined findings that could be grouped in the general intervention category of "political governance". While these are not interventions in the traditional sense of the term, the political nature of terrorism broadens related responses to a wide variety of arenas. For example, Barros (2003) analyzed the effects of having a Socialist party in power (which he

describes as the more intolerant and harsher party against rogue political groups) in Spain on the effects of ETA terrorism, while Enders and Sandler (2000) examined the effect of the end of the cold war on terrorism time series.

The combination of findings in Figure 8 indicates an uncertainty about whether the existence of harsher parties on terrorism as well as the end of the Cold War may increase terrorism events. The weighted mean effect size of these findings was 6.16, although the left side of the confidence interval barely indicated a null effect (CI: {-0.46, 12.79}). However, the homogeneity test did not indicate an strongly significant chi-squared statistic. When running a fixed effects model of this category, it revealed that a statistically significant harmful effect of intolerance or the end of the cold war on terrorism (weighted mean = 4.23, CI: {2.56, 6.20}).

Also interesting when examining individual findings was that the harmful effect of both an intolerant party as well as the end of the cold war was reflected in more dangerous outcomes (assassinations, and events which led to individuals becoming wounded or dying) while these aspects of political governance reduced the likelihood of less serious, non-casualty events.

Figure 8. Political Governance: Socialist Party in Power or Post Cold War



6 Discussion

The combination of findings from the seven studies found to evaluate counter-terrorism measures using at least a moderately strong research design provides for a number of interesting discussion points across this literature. Most importantly, interventions commonly used, including military retaliation campaigns, the fortification of buildings, United Nations resolutions and other laws, as well as increasing the certainty and/or severity of punishment, may not be as promising as much of the non-evaluation terrorism literature often seems to suggest. In many cases, effects are not statistically discernible from a null effect, or can be harmful and increase the likelihood of certain types of terrorism.

Furthermore, different effects may occur depending on what outcome is being measured. Metal detectors "work" in reducing airplane hijackings, but as Cauley and Im (1988) and Enders and his colleagues have emphasized, there may be displacement or substitution effects leading to increases in other types of terrorism not involving aircraft. Military retaliations can also lead to increases in terrorism, although it appears that these increases may be short-term, less lethal threats and other terrorist activities that do not use substantial resources.

Perhaps what is equally (if not more) interesting is what we *didn't* find from our review. Only a small subset of interventions was analyzed across the seven studies, using one primary methodology (time series) and one general outcome type (terrorism incidents). However, when we examined the 94 studies which seemed remotely connected to some evaluation, a number of interventions were discussed, including:

- airport screening
- anti-terrorism products for personal use
- arrest
- assassinations
- blast resistant luggage
- building security
- detection devises for biological or chemical weapons
- diplomatic efforts
- educational support
- emergency response preparedness
- fortification of embassies
- gas masks
- hostage negotiation
- laws against terrorism
- medical antidotes for anthrax
- medical antidotes for nerve gas
- medical antidotes for other chemical or biological weapons
- medical antidotes for smallpox
- metal detectors
- preparedness for bioterrorist attack
- prosecution strategies
- psychological counseling

- psychological treatment for trauma or PTSD
- punishment and sentencing
- religious interventions
- seal/tamper proof devices
- social and economic responses
- UN Conventions and multilateral interventions
- use of force by one country to another
- use of media
- vaccinations

We suspect this list only partially represents the different types of counter-terrorism interventions that exist that can be evaluated.

The findings strongly confirm our initial speculation about the state of counter-terrorism strategies. There has been a proliferation of anti-terrorism programs and policies as well as massive increases in expenditures toward combating terrorism. Yet, we know almost nothing about the effectiveness of any of these programs. The seven studies we found, from a total of over 20,000 pieces of literature, focus on only a handful of specific interventions, and often involve the same researchers examining similar datasets, using similar methodologies. **Clearly, the most important policy recommendation to emerge from this review is that the wide-array of anti-terrorism policies need to be evaluated for effectiveness or at least be better informed by existing scientific evaluations. Related to this is the importance of funding and encouragement for scientific evaluations of counter-terrorism programs.**

The small amount of evaluation research in the area of terrorism may be due to a number of reasons. First, practical problems in either qualitatively or quantitatively studying terrorism are often thought to be significant obstacles to overcome. Not only are events thought of as "rare", but units of analysis (individuals or groups who employ terror violence or actual events/incidents of terrorism) may be difficult to locate and study. Unlike criminological research that contains large amounts of data from frequently occurring events, terrorism research relies on much scarcer occurrences, which makes analysis more challenging (in terms of identifying patterns). Furthermore, challenges that do appear with using crime data are accentuated in terrorism prevention programs. These include the difficulty in detecting intervention effects of major programs where treatment effects may be unclear, or alternative explanations are difficult to control for. For terrorism, threats are often not clearly defined so the impacts of the programs are almost impossible to assess.

Compounding this problem is the nature of the counter-terrorism enterprise itself, which, unlike much of current local policing, is shrouded in secrecy. Because of this, information about incidents and outcomes are difficult to collect for scientific purposes. There is also a real problem in establishing, for the data that are available, clear standards of accuracy and reliability, problems that have been exhaustively addressed in the reform of police statistical standards through UCR reporting and other standardization procedures. Further, in terms of evaluation, there is likely to be less willingness in this area to engage in experimental programs that might test differential program effects (a reluctance that was to a degree overcome in police evaluations, such as those performed by Sherman et al. (1992) in the area of domestic violence interventions). The evaluations, then, would have to be conducted ex post facto, as illustrated in the case studies that we report above. While these studies reveal some interesting and important

effects, their real impact is diminished through the inability to adjust program effects or increase the sensitivity of measurement in an ongoing assessment process.

It is certainly the case that researchers can extend the types of evaluations that would use existing data sets to test for the effects of major interventions, such as the Patriot Act. The availability of the ITERATE data base, the files available from MIPT, and the new data that will come on line through initiatives supported by DHS funding will encourage more detailed, albeit retrospective, analysis of terrorism. As we move forward, it is important that these data are improved through the addition of important contextual information that will help uncover the multitude of factors that influence these events.

In addition, there needs to be a new initiative that evolves from the intelligence community that would mirror the increased collaborations between researchers and police over the last few decades. It is clear from the extensive literature that has appeared in recent years, as documented above, that there is no real shortage of interest in examining terrorism. The challenge is to introduce the same rigorous standards of research that are applied in other areas of criminal justice and at the same time convince the intelligence community that we are trying to help. A willingness to engage in evaluation studies and to manage interventions in a way that allows for systematic assessment of effects is an important part of this new agenda. Improving data quality through the introduction of reporting standards and reliability measures would go a long way to enhancing these initiatives.

In terms of future research, the most productive area for evaluation may lie in examining both the effectiveness and effects of counter-terrorism programs. As already emphasized, connecting interventions with outcomes is central to successful evidence-based policy. However, with programs targeting rare events more specifically, there will be collateral consequences that also must be monitored. For example, aside from decreasing terrorism events, what is the effect of a counter-terrorism program on individuals' fear, civil liberties, lifestyles, or mental/physical health? How do programs affect public opinion toward a government, leader, nation, or even an ideology (for example, "democracy" or "Islam") or symbol ("the West")? How might strategies improve or worsen diplomatic relationships across nations? How does a shift in law enforcement resources and functions affect the relationship between police and minority groups? These questions are directly related to the effectiveness of counter-terrorism programs and are as important to study as concerns of event reduction.

7 Recommendations to Decision-Makers

In conclusion, we make the following recommendations to three types of decision makers who may take interest in the information provided by this review. These decision makers are government agencies (and their associated agents) who generate counter-terrorism policy, agencies who fund counter-terrorism research, and researchers involved in the study of terrorism.

7.1 TO GOVERNMENT AGENCIES GENERATING COUNTER-TERRORISM POLICY

1. Counter-terrorism policy needs to be rational, effective, and cause as little harm as necessary. In order to determine the effectiveness of counter-terrorism strategies, we need to evaluate whether the intervention is connected to a measurable, desired outcome. Surveying individuals about whether they think a policy works, or subjectively determining what a successful strategy "looks like" is neither scientific nor will it generate the necessary data/information to allow for a determination to be made regarding the effectiveness of a particular strategy.

2. Along the same lines, government agents need to pay attention to scientific research about counter-terrorism programs when deciding upon policy choices. Such attention can also serve as a reciprocal force to increase evaluation research by the academic community. It is an often-mistaken notion that research is far removed from practice; countless examples of the effective use of findings from evaluation in crime prevention have proven otherwise.

3. This need for evaluation research requires that scientists are welcomed into the counter-terrorism enterprise. This includes fostering better collaborations between government agencies and researchers. Indeed, researchers understand that much information associated with terrorism requires security protections, and government agencies should consider extending clearances to evaluation researchers. However, some of this secrecy is unwarranted or at least can be better facilitated. Many decades ago, police agencies also mistakenly believed that crime data could not be given to scientists to study. Luckily, many police agencies have overcome these fears, which have led to major improvements in their effectiveness. The interaction between social scientists and government has led to many social advances, and the area of terrorism is no exception.

7.2 TO GOVERNMENT AGENCIES FUNDING RESEARCH

1. Funding should be directed to the evaluation of existing counter-terrorism programs, rather than "process descriptions" which describe procedures or determine if procedures were carried out according to a plan.

2. Evaluations should be at least moderately rigorous in design quality. Using less rigorous evaluations will result in findings that are less reliable or believable, and may falsely exaggerate effects (see Weisburd et al., 2001). As Lum and Yang (2005) have found, government agencies which fund research have been able to influence the quality of research through the power of the purse. We recommend that agencies funding evaluation research on counter-terrorism strategies demand higher-quality evaluation designs.

3. Along the same lines, funding agencies should encourage the discovery of alternative or improved ways to gauge effectiveness of counter-terrorism programs. For example, could experimental and other types of quasi-experimental designs (aside from time-series) be used to evaluate some programs (perhaps those used to respond and manage terrorism and terrorism-related byproducts)? Not only is funding needed for the evaluation research itself, but also in endeavors to discover how to evaluate terrorism outcomes and interventions.

7.3 TO RESEARCHERS STUDYING TERRORISM

1. Only a small fraction of terrorism research appears to be empirical in nature. We need more empirical research on terrorism, perhaps that goes beyond just examining numbers of events over time. More specifically, evaluations of interventions are necessary, such as those listed in the discussion of this review.

2. Not only is more evaluation research needed, but researchers need to also explore ways in which terrorism strategies can be analyzed and different types of methodologies which may be useful.

3. Evaluation research can serve as a moderating and rational effect on rash policy responses based on moral panic and fear. These types of policies often lead to other social negatives, including the violation of personal and human rights as well as further humiliation that can lead to more terrorism. Scientists can provide this moderating effect with more research that evaluates both the outcome effectiveness as well as the social, political, economic, or psychological effects of these interventions.

8 Plans for Updating the Review

We will update this review beginning in January of 2007. In the update, we also plan to examine the possibility of including studies which use other methods that are not normally included in methodological rigor scales, but that have emerged as possible sources of evaluation in terrorism studies. In particular, country and case studies will be examined to determine if they are appropriate to use to determine "what works" in counter-terrorism strategies. Additionally, we hope to overcome the current limitation of excluding non-English studies. Finally, reviewers suggested including other terms in our search, including "separatist", "insurgents", "outcome", "consequence" and/or "result" and also add the Australian database CINCH (Computerized Index to National Criminological Holdings).[19]

[19] Thanks to Peter Grabosky for these suggestions.

9 Acknowledgements

We would like to acknowledge with gratitude the financial support of the Campbell Collaboration, and the Jerry Lee Center of Criminology at the University of Pennsylvania for supporting the preparation of the protocol and review. The original funding for the bibliographical review was provided by the Office of the Provost, Rutgers Newark and the Rutgers School of Criminal Justice to the Center for the Study of Public Security at Rutgers Newark. In particular, we are indebted to David Wilson for his methodological assistance and comments. Special thanks to Louise Stanton, Michelle Grillo, Lauren Giordano, Deborah Vegh and Sue-Ming Yang for their research assistance. Helpful comments were also provided by Peter Grabosky and the anonymous reviewers.

10 References

Abadie, A. and Gardeazabal. J. (2001). *The economic costs of conflict: A case control study for the Basque Country*. NEBR Working Paper 8478. Cambridge, MA.

Barros, C. P. (2003). An intervention analysis of terrorism: The Spanish Eta case. *Defence and Peace Economics, 14*(6), 401-412.

Bisson, J. and Andrew, M.. (2005). Psychological treatment of post-traumatic stress disorder (PTSD). *The Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD003388.pub2. DOI: 10.1002/14651858.CD003388.pub2.

Boruch, Robert F., Snyder, B., and DeMoya, D. (2000). The importance of randomized field trials. *Crime and Delinquency, 46,* 156-80.

Bozzette, Samuel A., Boer, R., Bhatnagar, V., Brower, J. L., Keeler, E. B., Morton, S. C. et al. (2003). A model for smallpox-vaccination policy. *New England Journal of Medicine, 348*(5), 416-425.

Brophy-Baermann, B. and Conybeare, J. A.. (1994). Retaliating against terrorism: Rational expectations and the optimality of rules versus discretion. *American Journal of Political Science, 38*(1), 196-210.

Cauley, J. and Im, E. (1988). Intervention policy analysis of skyjackings and other terrorist incidents. *The American Economic Review, 78*(2), 27-31.

Chauncey, R. (1975) Deterrence: Certainty, deterrence, and skyjacking. *Criminology, 12*(4), 447- 473.

Cohen, S. (1972). *Folk devils and moral panics: The creation of the mods and rockers*. Oxford: Blackwell.

Congressional Budget Office. (2002). *The budget and economic outlook: Fiscal years 2003 to 2012.* Washington, DC: Congress of the United States.

Congressional Budget Office. (2005). *The budget and economic outlook: Fiscal years 2006 to 2015.* Washington, DC: Congress of the United States.

Cook, T. and Campbell, D. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

Crenshaw, M. (1992). Current research on terrorism: The academic perspective. *Studies in Conflict and Terrorism, 15*(1), 1-11.

Cullen, F. T. and Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In Horney, J. (Ed.), *Policies, processes, and decisions of the criminal justice system: Criminal justice 3*. Washington, DC: National Institute of Justice, U.S. Department of Justice.

Davies, H.O., Nutley, S. and Smith, P. C. (2000). *What works: Evidence-based policy and practice in public services.* London: Policy Press.

Enders, W. and Sandler, T. (1993). The effectiveness of antiterrorism policies: A vector-autoregression-intervention analysis. *The American Political Science Review, 87*(4), 829-844.

Enders, W. and Sandler, T. (2000). Is transnational terrorism becoming more threatening? *Journal of Conflict Resolution, 44*, 307-332.

Enders, W. and Sandler, T. (2002). Patterns of transnational terrorism, 1970-1999: Alternative time-series estimates. *International Studies Quarterly, 46*(2), 145-165.

Enders, W., Sandler, T., and Cauley, J. (1990). UN conventions, terrorism, and retaliation in the fight against terrorism: An econometric evaluation. *Terrorism and Political Violence, 2*(1), 83.

Farrington, D. (2003). Methodological quality standards for evaluation research. *The Annals of the American Academy of Political and Social Science, 587*(1), 49-68.

Farrington, D. and Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *Annals of the American Academy of Political and Social Science, 58*, 35-49.

Gillespie, K., Duffy, M., Hackmann, A. and Clark, D. (2002). Community based cognitive therapy in the treatment of posttraumatic stress disorder following the Omagh bomb behaviour. *Research and Therapy, 40*, 345-357.

Guinnessy, J., and Dawson. P. (2002). Terrorism drives Bush R and D money to defense and NIJ; The science funding flat in fiscal 2003. *Physics Today, 55*, 30.

Halkides, M. (1995). How not to study terrorism. *Peace Review 7*, 253-260.

Halloran, M. El., Longini Jr., I. M., Nizham, A. and Yang, Y. (2002). Containing bioterrorist smallpox. *Science, 298*, 1428-1432.

Hoffman, B. (1992). Current research on terrorism and low-intensity conflict. *Studies in Conflict and Terrorism, 15*, 25-37.

Issues in Science and Technology (Report). (2002). Federal R and D in FY 2002 will have biggest percentage gain in 20 years. *Issues in Science and Technology, 18*, 23.

Jefferson T., Demicheli, V., Deeks J., Graves, P., Pratt, M., and Rivetti, D.. (1998). Vaccines for preventing anthrax. *The Cochrane Database of Systematic Reviews*, Issue 1. Art. No.: CD000975. DOI: 10.1002/14651858.CD000975.

Johnston, R. G., Garcia, A. R. E., and Pacheco, A. (2002). The efficacy of tamper indicating devices. *Journal of Homeland Security, April.* Los Alamos National Laboratory, Vulnerability Assessment Team. Available online at: http://www.homelandsecurity.org/journal/articles/displayArticle.asp?article=50.

Lallement, G., Clarencon, D., Masqueliez, C., et al. (1998). Never agent poisoning in primates: Antiletal, anti-epileptic and neuroprotective effects of GK-11. *Archives of Toxicology, 72*, 84-92.

Landes, W.M. (1978). An economic study of U.S. aircraft hijackings, 1961-1976. *Journal of Law and Economics 21*, 1-31.

LeVine, V. T. and Salert, B. A. (1996). Does a coercive official response deter terrorism? The case of the PLO. *Terrorism and Political Violence, 8*(1), 22-49.

Lipsey, M. and Wilson, D. (2001). *Practical meta-analysis.* Applied Social Research Methods Series 49. Thousand Oaks: Sage.

Lum, C. and Yang, S. (2005). Why do evaluation researchers in crime and justice choose non-experimental methods? *Experimental Criminology, 1*, 191-213.

Macilwain, C. (2002). Bush goes to war as budget boosts R and D. *Nature, 415*, 564.

MacKenzie, D. (2000). Evidence-based corrections: Identifying what works. *Crime and Delinquency, 46*, 457-71.

Martz, H. and Johnson, M. (1987). Risk analysis of terrorist attack. *Risk Analysis, 7*(1), 35-47.

McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. *Annals of the American Academy of Political and Social Science* 587:16-30.

Merari, A. (1991). Academic research and government policy on terrorism. In C. McCauley (Ed.), *Terrorism research and public policy*. London: Frank Cass.

Mickolus, E. (1980). *Transnational terrorism: A chronology of events 1968 – 1979.* Westport: Greenwood.

Mickolus, E. (1982). *International terrorism: Attributes of terrorist events, 1968 – 1977 (ITERATE 2).* Ann Arbor: Inter-University Consortium for Political and Social Research.

Mickolus, E., Sandler, T. and Murdock, J. (1989). *International terrorism in the 1980s: A chronology of events. 2 Volumes*. Ames: Iowa State University Press.

Mickolus, E., Sandler, T., Murdock, J. and Fleming, P. (1989). *International terrorism: Attributes of terrorist events 1978 – 87 (ITERATE 3).* Dunn Loring: Vinyard Software.

Mickolus, E., Sandler, T., Murdock, J. and Fleming, P. (1993). *International terrorism: Attributes of terrorist events 1988-91 (ITERATE 3).* Dunn Loring: Vinyard Software.

Miller, R. (1988). The literature of terrorism. *Terrorism, 11*, 63-87.

Nutley, S. and Davies, H. O. (1999). The fall and rise of evidence in criminal justice. *Public Money and Management, 19*, 47.

Prunckun, H. and Mohr, P. (1997). Military deterrence of international terrorism: An evaluation of Operation El Dorado Canyon. *Studies in Conflict and Terrorism, 20*, 267-280.

Quinn, C., Semenova, V., Elie, C. M., et al. (2002). Specific, sensitive, and quantitative enzyme-linked immunosorbent assay for human immunoglobulin G antibodies to anthrax toxin protective antigen. *Emerging Infectious Diseases, 8*(10), 1103-1110.

Romano, T. (1984). *Terrorism: An analysis of the literature.* Dissertation, Fordham University, Department of Sociology, Criminology and Penology.

Rose, S., Bisson, J., Churchill, R., and Wessely, S. (2002). Psychological debriefing for preventing post traumatic stress disorder (PTSD). *The Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD000560. DOI: 10.1002/14651858.CD000560.

Schmid, A. (1983). *Political terrorism: A research guide to concepts, theories, databases, and literature.* New Brunswick: Transaction Books.

Schmid, A. P. and Jongman, A.J. (1988). *Political terrorism: A new guide to actors, authors, concepts, databases, theories and literature*. Amsterdam: North-Holland Publishing Company.

Shadish, W., Cook, T. and Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inferences.* Boston: Houghton-Mifflin.

Sherman, L. W. (1992). *Policing domestic violence: Experiments and dilemmas.* New York: Free Press.

Sherman, L. W. (1998). *Evidence-based policing*. Second Invitational Lecture on Ideas in Policing. Washington, DC: Police Foundation.

Sherman, L.W., Farrington, D.P., Welsh, B.C. and MacKenzie, D.L. (Eds.). (2002). *Evidence based crime prevention.* London: Routledge.

Sherman, L.W., Gottfredson, D., MacKenzie, D.L., Eck, J., Reuter, P. and Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising: A report to the United States Congress.* Washington, DC: National Institute of Justice.

Silke, A. (ed.). (2004). *Research on terrorism: Trends, achievements and failures.* New York: Taylor and Francis.

Silke, A. (2005). Fire of Iolaus: The role of state counter-measures in causing terrorism and what needs to be done. In T. Bjorgo (Ed.), *Root causes of terrorism: Myths, reality and ways forward*. Oxford, UK: Routledge.

Smith, Brent L. and Orvis, Gregory P. (1993). America's Response to Terrorism: An Empirical Analysis of Federal Intervention Strategies During the 1980's. *Justice Quarterly, 10*(4): 661-681.

Smith, B. L., Damphousse, K. R., Jackson, F. and Sellers, A. (2002). The Prosecution and punishment of international terrorists in federal courts: 1980-1998. *Criminology and Public Policy, 1*(3), 311-338.

Stein, D.J., Zungu-Dirwayi, N., van der Linden, GJH., and Seedat, S. (2000). Pharmacotherapy for post traumatic stress disorder (PTSD). *The Cochrane Database of Systematic Reviews*, Issue 4. Art. No.: CD002795. DOI: 10.1002/14651858.CD002795.

U.S. House of Representatives. (2004). *Effective strategies against terrorism: Hearing before the Subcommittee on National Security, Emerging Threats and International Relations of the Committee on Government Reform.* One Hundred Eighth Congress, Second Session, February 3. Available online at: http://www.mipt.org/pdf/Effective-Strategies-Against-Terrorism.pdf.

Walter, E.V. (1969). *Terror and resistance: A study of political violence with case studies of some primitive African communities.* New York: Oxford University Press.

Weisburd, D., Lum, C. and Petrosino, A. (2001). Does research design affect study outcomes in Criminal Justice? *The Annals of the American Academy of Political and Social Science, 578*, 50-70.

Weisburd, D., Petrosino, A. and Lum, C. (Eds.). (2003). Assessing systematic evidence in crime and justice: Methodological concerns and empirical outcomes (Preface). *The Annals of the American Academy of Political and Social Science, 587*, 6-14.

Wilkinson, P. (1977). *Terrorism and the liberal state.* New York: John Wiley and Sons, Inc.

Zussman, A. and Zussman, N. (2005a). *Targeting killings: Evaluating the effectiveness of a counterterrorism policy.* Discussion Paper. Bank of Israel, Research Department.

Zussman, A. and Zussman, N. (2005b). *Assassinations: Evaluating the effectiveness of a counterrorism policy using stock market data*. Cornell University, Dept. of Economics.

11 Appendices

APPENDIX A: Organizations Conducting Terrorism Research

ORGANIZATION	URL
Adolescents in Political Violence Project (University of Tennessee)	http://cfs.utk.edu/f_s/barber.html
ANSER Institute for Homeland Security	http://www.homelandsecurity.org
Aon Corporation	http://www.aon.com/
Belfer Center for Science and International Affairs (Harvard University)	http://bcsia.ksg.harvard.edu/
Bioterrorism Preparedness Office (Center for Disease Control)	http://www.cdc.gov/
Board of Neuroscience and Behavioral Health (National Academy of Science)	http://www.iom.edu
Homeland Security Project (Brookings Inst.)	http://www.brookings.edu/fp/research/projects/homeland/homeland.htm
Canadian Security Intelligence Service (Canadian Government)	http://www.csis-scrs.gc.ca/eng/menu/welcome_e.html
Center for Defense and International Security Studies	http://www.cdiss.org/terror.htm
Center for Biosecurity and Public Health Preparedness (University of Texas – Houston)	http://www.sph.uth.tmc.edu/cbphp
Center for Civilian Biodefense Strategies (Johns Hopkins University)	http://www.hopkins-biodefense.org
Center for Contemporary Conflict (US Navy)	http://www.ccc.nps.navy.mil/
Center for Defense Information - Terrorism Project	http://www.cdi.org
Center for Democracy and Technology	http://www.cdt.org/
Center for International Security and Cooperation (Stanford University)	http://cisac.stanford.edu
Center for Non-Proliferation Studies (Monterey Institute of International Studies)	http://cns.miis.edu/
Center for Peace and Security (Georgetown University)	http://cpass.georgetown.edu/
Center for Public Health and Disasters (UCLA)	http://www.cphd.ucla.edu
Center for Technology and National Security Policy (US Department of Defense)	http://www.ndu.edu/ctnsp/index.html
Center for Terrorism Preparedness (University of Findlay)	http://www.nceem.org/terrorism/default.asp
Center for the Prevention of Genocide (Improve the World International)	http://www.genocideprevention.org/index.htm

Center for the Study of Bioterrorism and Emerging Infections (St. Louis University)	http://bioterrorism.slu.edu/
Center for the Study of Public Security (Rutgers University)	http://www.andromeda.rutgers.edu/~rcst/home.html
Center for the Study of Terrorism and Political Violence (University of St. Andrews)	http://www.st-andrews.ac.uk/academic/intrel/research/cstpv/
Center on Terrorism and Irregular Warfare (US Navy)	http://www.nps.navy.mil/ctiw/
Center on Terrorism and Public Health (Florida State College of Medicine)	http://www.med.fsu.edu/healthaffairs/ctph/default.asp
Center on Terrorism and Public Safety (John Jay College of Criminal Justice)	http://www.centeronterrorism.org
Centers for Public Health Preparedness (Center for Disease Control)	http://www.phppo.cdc.gov/owpp/CPHPLocations.asp
CERT Coordination Center (Carnegie Mellon University)	http://www.cert.org/
Chemical and Biological Arms Control Institute	http://www.cbaci.org/
Chemical and Biological Arms Control Program (Federation of American Scientists)	http://www.fas.org/bwc/index.htm
Chemical and Biological Defense Information Analysis Center (Battelle Memorial Institute)	http://www.cbiac.apgea.army.mil/
Columbia University World Trade Center Archive Project	http://www.columbia.edu/cu/lweb/news/libraries/2001-10-30.wtc_archives.html
Command and Control Research Program (Department of Defense)	http://www.dodccrp.org/
Conflict Archive on the Internet (University of Ulster)	http://cain.ulster.ac.uk/
Congressional Research Service (Library of Congress)	http://www.loc.gov/crsinfo/whatscrs.html#about
Council on Foreign Relations	http://www.cfr.org/
Counterterrorism Office (Department of State)	http://www.state.gov/s/ct/
Critical Incident Analysis Group (University of Virginia)	http://www.healthsystem.virginia.edu/ciag
Dart Center for Journalism and Trauma (University of Washington)	http://www.dartcenter.org/
Defense Advanced Research Projects Agency (Department of Defense)	http://www.darpa.mil/
Defense Information Systems Agency (Department of Defense)	http://www.disa.mil/
Defense Intelligence Agency (Department of Defense)	http://www.dia.mil/
Emergency Response and Research Institute	http://www.emergency.com
Facts on File, Inc (Ferguson Publishing Co.)	http://www.factsonfile.com/
Federal Research Division-Terrorism and Crime Studies (Library of Congress)	http://lcweb.loc.gov/rr/frd/
Financial Action Task Force on Money Laundering	http://www1.oecd.org/fatf/index.htm
Foreign Policy Research Institute	http://www.fpri.org/
Foundation for the Defense of Democracies	http://www.defenddemocracy.org/

National Security Archive - George Washington University	http://www.gwu.edu/~nsarchiv/
Government Accounting Office- Special Collections- Terrorism	http://www.gao.gov/terrorism.html
Henry L. Stimson Center	http://www.stimson.org
Human Rights Watch	http://www.hrw.org/
Institute for Security Technology Studies (Dartmouth College)	http://www.ists.dartmouth.edu/
International Center for Terrorism Studies (Potomac Institute for Policy Studies)	http://www.potomacinstitute.org/academic/icts.cfm
International Critical Incident Stress Foundation	http://www.icisf.org/
International Policy Institute for Counter-Terrorism	http://www.ict.org.il/
International Society for Traumatic Stress Studies	www.istss.org
Jaffe Center for Strategic Studies at Tel Aviv University	http://www.tau.ac.il/jcss/about.html
Library of Congress September 11 Archive	http://september11.archive.org/
The Mackenzie Institute	http://www.mackenzieinstitute.com/
Matthew B. Ridgway Center for International Security Studies (University of Pittsburgh)	http://www.gspia.pitt.edu/ridgway/
Mitretek Systems	http://www.mitretek.org/home.nsf
Narcoterror.org	http://www.narcoterror.org/
National Association of Insurance Commissioners	http://www.naic.org/
National Geospatial Intelligence Agency (formerly National Imagery and Mapping Agency, Department of Defense)	http://www.nima.mil/
National Security Agency	http://www.nsa.gov/
National Technical Information Service (Department of Commerce)	http://www.ntis.gov
Nuclear Control Institute	http://www.nci.org/
US Customs and Border Protection (DHS)	http://www.customs.ustreas.gov/
Office of Domestic Finance—Terrorism Risk Insurance Program (Department of Treasury)	http://www.treasury.gov/offices/domestic-finance/financial-institution/terrorism-insurance/
Office of Foreign Asset Control (Department of Treasury)	http://www.ustreas.gov/offices/enforcement/ofac/
Oklahoma City National Memorial Institute for the Prevention of Terrorism	http://www.mipt.org/
Pacific Northwest National Laboratory (Department of Energy)	http://www.pnl.gov/
PILOTS catalog – Dartmouth College Library	http://www.dartmouth.edu
Pinkerton Global Intelligence Services	http://pgis.pinkertons.com/

Political Risk and Trade Credit Group (MMC Enterprise Risk Group)	http://www.mmcer.com/
Public Health Practice Program (Center for Disease Control)	http://www.phppo.cdc.gov
RAND	http://www.rand.org/
SAPRA India	http://www.subcontinent.com/sapra.html
South Asia Terrorism Portal (Institute for Conflict Management)	http://www.satp.org/
SSAF International Ltd.	http://www.ssafprotection.com/
Technical Support Working Group (Department of State)	http://tswg.gov
Technology and Public Policy Program (Center for Strategic and International Studies)	http://www.csis.org/tech/index.htm
Global Project Against Terrorism (Technical Assistance programme of the Terrorism Prevention Branch, United Nations)	http://www.odccp.org/odccp/terrorism.html
Terrorism Research Center	http://www.homelandsecurity.com/
The Cato Institute	http://www.cato.org/
The New York Times	http://www.nytimes.com/
The Washington Post	http://www.washingtonpost.com/
Transnational Threats Initiative (Center for Strategic and International Studies)	http://www.csis.org/tnt/
United States Institute of Peace	http://www.usip.org/
US Department of Agriculture – Food Safety and Inspection Service; Agricultural Research Service	http://www.fsis.usda.gov/; http://www.ars.usda.gov/
US Food and Drug Administration—Center for Biologics Evaluation and Research; and Center for Food Safety and Applied Nutrition	http://www.fda.gov/cber/index.html; http://vm.cfsan.fda.gov/list.html
US Secret Service	http://www.secretservice.gov/index.shtml
US Air Force Counter Proliferation Center	http://www.au.af.mil/au/awc/awcgate/awc-cps.htm http://c21.maxwell.af.mil/
Versar Inc.	http://www.versar.com/
Wade Financial Group	http://www.terrorism-insurance-risk-management.com/
William R. Nelson Institute for Public Affairs (James Madison University)	http://www.jmu.edu/orgs/wrni/

APPENDIX B: Data Sources Available

ORGANIZATION	DATA DESCRIPTION
Avalon Project, Yale Law School	http://www.yale.edu/lawweb/avalon/terrorism/terror.htm documents pre-18th century
CDISS Terrorism Program - Center for Defense and International Security Studies	Terrorist incidents 1945 to 1998
Center for the Prevention of Genocide	Country reports online
Center for Non-Proliferation Studies	Weapons of mass destruction database
Center for the Study of Terrorism and Political Violence- University of St. Andrews	Terrorism incident database
Center on Terrorism and Irregular Warfare	Government databases. Suicide bombers.
Chemical and Biological Defense Information Analysis Center	CBIAC Bibliographic Database
Columbia University World Trade Center Archive Project	Archive data on September 11th
Conflict Archive on the Internet	Databases on Northern Ireland conflict
Counterterrorism Office- US Department of State	Patterns and trends of terrorism, chronology online
Department of Psychology-Political Violence Program- Tel Aviv University	Ariel Merari database on terrorism incidents
Emergency Response and Research Institute	Privately-held databases on terrorist personnel, terrorism methods, terrorist incidents, and the implications thereof
Federal Research Division-Terrorism Studies	Terrorism databases
Federation of American Scientists (FAS), Intelligence Resource Program, "Liberation Movements, Terrorist Organizations, Substance Cartels, and other Para-State Organizations."	http://www.fas.org/irp/world/para/index.html This directory of para-states is not a list of terrorist organizations, and is not constructed to supplement or complement the list of terrorist organizations of the US Department of State. The guide intentionally casts a wide net, and includes both the nasty and nice.
George Washington University	September 11 Source Books

Henry L. Stimson Center	Unconventional weapons, chemical and biological terrorism
International Center for Political Violence and Terrorism Research (ICPVTR)	http://www.ntu.edu.sg/idss/research_03a.htm The ICPVTR database focuses on the Asia Pacific region, especially Southeast Asia. No other details are available and it is not yet available to the public.
International Policy Institute for Counter-Terrorism (ICT) – Israel	Houses six databases on terrorists groups, incidents, activity
Library of Congress	Archive of digital materials up after 9/11
University of Michigan Documents Center "America's War Against Terrorism—World Trade Center/Pentagon Terrorism and the Aftermath"—	http://www.lib.umich.edu/govdocs/usterror.html
National Association of Insurance Commissioners	Insurance related databases such as claims from attacks and other criminal liabilities
The New York Times	News archives
Oklahoma City National Memorial Institute for the Prevention of Terrorism	RAND Terrorism Chronology Database, RAND-MIPT Terrorism Incident Database, and MIPT Indictment Database
PILOTS catalog	Published International Literature On Traumatic Stress
Pinkerton Global Intelligence Services	Incident Database (currently housed and researched at the University of Maryland, Department of Criminology under Professor Gary LaFree)
RAND	RAND Terrorism Incident Database
SSAF International Ltd	Multiple databases
Technical Support Working Group (US Department of State)	
Terrorism Prevention Branch- UN Office on Drugs and Crime	
Center for Non-Proliferation Studies	Terrorist Group Profiles
Terrorism Research Center	Terrorist group profiles, significant events, terrorist attacks, counter-terrorism group profiles, and country profiles.
Wade Financial Group	Insurance related information
The Washington Post	News archives

Notes