

Is counter-terrorism policy evidence-based? What works, what harms, and what is unknown

Cynthia Lum, Leslie W. Kennedy* and Alison Sherley*
George Mason University and * Rutgers University

Is counter-terrorism policy evidence-based? What works, what harms, and what is unknown. One of the central concerns surrounding counter-terrorism interventions today, given the attention and money spent on them, is whether such interventions are effective. To explore this issue, we conducted a general review of terrorism literature as well as a Campbell systematic review on counter-terrorism strategies. In this article, we summarize some of our findings from these works. Overall, we found an almost complete absence of evaluation research on counter-terrorism strategies and conclude that counter-terrorism policy is not evidence-based. The findings of this review 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.

¿Está basada la política antiterrorista en la evidencia? Lo que es efectivo, lo que es perjudicial, y lo desconocido. Una de las mayores preocupaciones de las políticas antiterroristas es su efectividad, dado el gasto enorme que suponen. Para explorar este punto llevamos a cabo una revisión sobre la literatura en general del terrorismo, así como una revisión Campbell sobre las estrategias antiterroristas. En conjunto hallamos casi una ausencia completa de evidencia sobre estos programas, y concluimos que la política antiterrorista no está basada en la evidencia. Esta revisión enfatiza la necesidad de que los políticos, los investigadores y los responsables de la gestión de estas políticas acentúen la necesidad de promover la evaluación de la efectividad de esos programas.

Recent events involving the use of violence and terror, most notably the attacks in the United States in 2001, have ignited major increases in personal, commercial, and governmental expenditures and attention on counter-terrorism strategies (Congressional Budget Office, 2002; 2005; Guinnessy & Dawson, 2002; Issues in Science and Technology, 2002; Lum et al., 2006[a], 2006[b]; Macilwain, 2002; Silke, 2004). For example, in 2005, the United States Congressional Budget Office estimated that U.S. defense spending for «Appropriations for Combating Terrorism and Protecting Critical Infrastructure» have increased more than ten-fold between 1998 and 2004, from US\$7.2 to \$88.1 billion (Congressional Budget Office, 2002; 2005) while non-defense funding for homeland security is estimated to have risen from \$9 billion in 2000 to \$32 billion in 2005.¹ Many other countries around the world have also experienced similar dramatic spending increases in an attempt to counter terrorism.

These expenditures have included a wide range of efforts, diverse in their goals, objectives, and ideological perspectives.

Some of these efforts consist of more traditional law enforcement approaches such as arrest, offender targeting, investigation strategies, or the expansion of police powers through the creation of new laws. Other efforts might be geared toward victims and can include treatment for post-traumatic stress disorder, improving the life-chances of individuals who suffer from wounds inflicted by explosives, developing antidotes for biological agents, or improving the emergency responsiveness of hospitals. Yet even other efforts may be political or social; strategies, tactics and counter-terrorism «programs» may include military action, third party negotiations, economic aid or sanctions, international resolutions, or efforts to influence the media. And, some programs are intended to reduce the opportunities for terrorism by hardening potential targets, such as increasing security screening at airports, placing barricades around buildings, or improving security protection for diplomats.

These remarkable trends in counter-terrorism spending and proliferation have led evaluation researchers, practitioners, and policy makers to question not only the effectiveness of these strategies, but how one might judge their effectiveness. Are the outcomes asserted (reductions in terrorism, terror-related risk, and harm caused by terrorism) indeed connected to these programs and do counter-terrorism interventions fulfill their promise? Furthermore, is counter-terrorism policy based in evidence of effectiveness or in something else (for example, personal opinions, political ideologies, or unscientific studies)?

The concept and movement of evidence-based social policy is grounded upon the assertion that choices to implement

¹ See <http://www.house.gov/budget/>.

intervention programs should be determined by what is known to be effective from conclusions drawn through scientific study and that currently implemented policies should also be subjected to scientific evaluations of effectiveness (Cullen & Gendreau, 2000; Davies et al., 2000; MacKenzie, 2000; Nutley & Davies, 1999; Petrosino et al., 2001; Sherman, 1998; Sherman et al., 2002; Weisburd et al., 2003; Welsh & Farrington, 2002). In both the medical and social sciences, the evidence-based policy movement has led to a call for more scientific evaluations of interventions using rigorous methods as well as meta-analyses and systematic reviews to summarize multiple evaluations for similar interventions that may vary in methodological quality and study design. Evidence-based interventions, therefore, are ones which have been supported by information and research, not opinions or political ideologies.

Given this, is counter-terrorism policy evidence-based? To examine this question, we began a general, comprehensive review of terrorism research, culminating in a Campbell Collaboration² systematic review published on the Campbell Website (see Lum et al., 2006[a]) and in an abridged version in the *Journal of Experimental Criminology* (see Lum et al., 2006[b]). The goal was to determine what is known about the effectiveness of counter-terrorism efforts, as well as where knowledge is lacking, to create a more informed policy and research agenda for evidence-based counter-terrorism approaches. Our major findings from these works were that only a small percentage of empirical studies of terrorism exist and there is an almost complete absence of evaluation research on counter-terrorism strategies. This is startling given the enormous increases in the development and use of counter-terrorism programs, as well as spending on counter-terrorism activity. Even more disconcerting was the nature of the evaluations we did find; some programs were shown to either have no discernible effect on terrorism or lead to *increases* in terrorism.

We summarize some of those findings here as an illustration of how counter-terrorism policy needs to be more evidence-based. This essay also suggests recommendations for policy makers, evaluations researchers, and funding agencies concerning the incorporation of empirical findings into the assessment of the effectiveness of counter-terrorism programs. Overall, we are left with the conclusion that counter-terrorism policy is indeed not evidence-based and steps should be taken to make it more so.

Background

We began our inquiry with a general overview of terrorism research. A comprehensive review of terrorism literature has not been attempted since September 11th (for past reviews see Halkides, 1995; Hoffman, 1992; Miller, 1988; Romano, 1984; Schmid & Jongman, 1988), and we expected this significant event to have had a major effect on the state of terrorism research. And, unlike some prevention programs in psychology, criminology or social work, we anticipated that strategies to counter terrorism would cut across disciplines and include programs not traditionally seen in the context of «crime prevention» (for example, the use of war or economic sanctions in an attempt to reduce the problem). To conduct this general overview of terrorism research, we examined all articles in published,

unpublished, peer-reviewed, non peer-reviewed, academic and non-academic sources which mentioned terms related to terrorism and political violence (for the purposes of this general review, we excluded books and some government documents, although these were included in the Campbell review). We conducted this search across seventeen separate literary databases,³ many of which extend back to research conducted since the early 1960s and which span multiple fields including medicine, criminology, psychology, political science, social work, sociology, education, and other physical sciences.

This preliminary search yielded over 14,000 records. However, the most unique finding compared to past literature reviews was of the distribution of the publication date of these studies. As Figure 1 indicates, among the works located, approximately 54% were published in 2001 and 2002.⁴ This is an exponential increase in writing on the subject, a trend which was not seen even after the Oklahoma City bombing or the Sarin gas attack on the Tokyo subway system in 1995, or the World Trade Center bombing in 1993. When examining articles from peer-reviewed sources, this same proportion was found. No other significant terrorist event has been followed by this much research interest on terrorism.

Evaluation research is a subset of empirical research more generally, and we also sought to understand to what extent terrorism research was based on empirical analyses (whether qualitative or quantitative). To gain a sense of the extent of empirical work on terrorism, we read each abstract of the peer-reviewed articles to see whether empirical analysis had been conducted. To be as inclusive as possible, we used a very broad definition of the term «empirical» – studies which indicated that any analysis (either quantitative or qualitative) had been conducted on terrorism information. Thought pieces, on the other hand, were articles where authors discussed an issue theoretically or offered an opinion, while case studies (as denoted by the author) examined a particular situation from a (usually) historical approach.

From this categorization, we found that only approximately 3% of the articles from peer-reviewed sources appeared to be based on some form of empirical analysis. Approximately 1% could be categorized as case studies and the rest (96%) were thought pieces. This rough estimate was extremely telling. 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 counter-terrorism strategies. More generally, it spoke to the state of terrorism research – that despite the efforts of some researchers to push empirical work forward, the general state of terrorism research lacks an empirical evidence base.

³ The databases used were Academic Search Premier, ArticleFirst (OCLC), Contemporary Women's Issues, 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/databasestrfold.pdf>.

⁴ This general review was initiated by the authors in 2003; hence, only literature up to the end of 2002 is represented.

² See <http://www.campbellcollaboration.org/>.

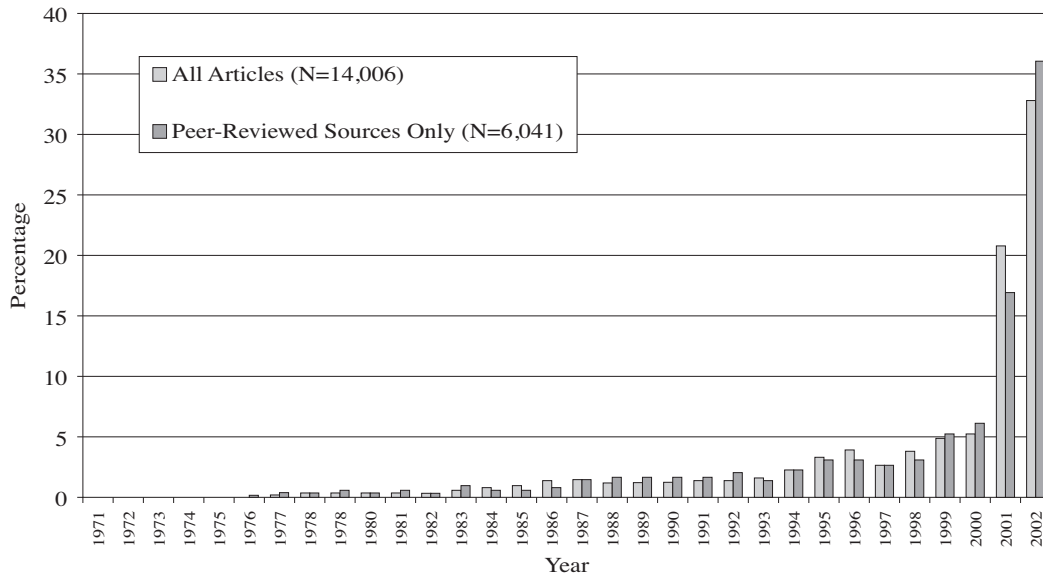


Figure 1. Yearly Distribution of Terrorism Publications as a Percentage of Total Publications

Given this lack of empirical work, as well as the very recent increase in terrorism studies interest, what then were writers discussing? To gain a better understanding of the subject matters of this research, we took all articles from peer-reviewed sources and categorized the studies by subject. Using this process, 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.

As table 1 indicates, 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, the sociology of terrorism (causes, motivations, explanations, definitions) and its impact were common topics.

	Peer- reviewed sources (N= 4,458 ^a)	Empirical studies 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

When examining those articles preliminarily deemed to be based on the analysis of empirical information (where we anticipated finding evaluation research), the findings are both encouraging and discouraging. A quarter of the empirical work appeared to be conducted on victimology and coping mechanisms (in particular, response and management strategies), a subject seemingly relevant to our search for evaluation studies on counter-terrorism programs. The next largest categories – the causes and sociology of terrorism as well as public attitudes to terrorism do not necessarily speak to the evaluation of counter-terrorism measures. Additionally, while political response pieces makes up 9.5% of the literature, only 1.9% of the empirical literature seems to empirically analyze those responses. Thus, there appears to be an overrepresentation of empirical literature on the explanations, causes and sociology of terrorism, rather than on programs designed to combat it. Overall, the empirically-based studies did not seem to emphasize evaluations.

These preliminary findings regarding the general state of terrorism research are revealing. 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 recent proliferation of counter-terrorism strategies, there is even more reason to evaluate these programs. Yet, there is a dearth of empirical research on counter-terrorism interventions, and within the empirical research no clear emphasis on evaluations. This is not to say that the current literature is not useful. However, this literature does not evaluate the effectiveness of the vast majority of counter-terrorism strategies and therefore we know little about whether measures might be effective or harmful.

The Campbell Systematic Review

While the general review described above provided an overall state of empirical terrorism research, Campbell systematic reviews utilize a specific process of searching for evaluations of interventions which satisfy a threshold of methodological rigor to be included in making conclusions about the effectiveness (or ineffectiveness) of social interventions. The systematic review has been detailed and published elsewhere (see Lum et al., 2006[a], 2006[b]) and therefore, only a summary of the evidence base of counter-terrorism policy is provided here.

Our Campbell review first began with establishing an initial search criteria for finding evaluation research on counter-terrorism strategies. 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, value-driven, and cover a wide-range of topics, areas, and subject matters. Thus, we chose to be highly inclusive in our initial search for evaluations of counter-terrorism research. We examined all studies that mentioned terrorism, no matter how defined, then searched for evaluations of interventions that might occur at any stage of the terrorism process, including prevention, detection, management or response strategies, as well as accepted a wide variety of possible measurable outcomes, including actual events as well outcomes such as fear or physical or mental healing.

Using our broad definitional and search criteria, we located over 20,000 written pieces of information that included books, articles, government and technical reports, online documents,

unpublished materials, dissertations, policy briefs, and other sources written about terrorism (our first search which yielded 14,000 documents only included articles, and not books or other types of publications). Within this large literature, we found only 354 studies that seemed to even hint at conducting or even discussing an evaluation of a counter-terrorism program or intervention. When examining each of these studies in detail, we found that only seven (7) were scientific evaluations of a counter-terrorism program. These seven evaluations are summarized in Table 2. Although at first glance the other studies appeared to be relevant, they were not, as they often just described the process of a program but did not evaluate it, made claims that a program was effective without any empirical test, advocated that evaluations should be done without doing any evaluation, interviewed people as to their personal beliefs about whether programs were effective, or described criteria by which a program might be evaluated but did not carry out any evaluations. We also excluded methodologically weak studies that did not use scientifically accepted evaluation approaches.

As Table 2 indicates, each of these studies (with the exception of Brophy-Baermann & Conybeare, 1994) was comprised of multiple findings for different interventions, time periods, and/or outcomes. Researchers reported evaluation results for multiple interventions on the same or different outcomes within the same study – for example, study authors examined the effect of metal detectors on reducing skyjackings, as well as on reducing embassy attacks. Additionally, because the studies were interrupted time series, we discovered 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 that connected an intervention to a measurable outcome within the seven studies.

In the Campbell review, we grouped the 86 findings into six general categories of interventions: (1) interventions which increased detection at airports, including installing metal detectors and increasing security screening more generally; (2) interventions 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) 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; and (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).

Results

Table 3 summarizes our findings for the subset of the 86 findings. For each intervention category, we provide the number of findings that indicated the counter-terrorism program for each category «worked» (there was a statistically significant decline in the rate of terrorism after the intervention was implemented), had «no statistically discernible effect» (there was no statistically significant change in the level of terrorism after the intervention was implemented) or that the intervention «was harmful» (showed a statistically significant *increase* in terrorism events after the intervention was implemented).

<i>Table 2</i> Summary of the seven studies in the Campbell Review in chronological order	
Citation	Summary
Landes, W.M. (1978). An Economic Study of U.S. Aircraft Hijackings, 1961-1976. <i>Journal of Law and Economics</i> , 21, 1-31	Used United States Federal Aviation Administration data on skyjackings to examine 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
Cauley, J., & Im, E. (1988). Intervention Policy Analysis of Skyjackings and Other Terrorist Incidents. <i>The American Economic Review</i> , 78(2), 27-31	Used an interrupted time series analysis for terrorism incidents occurring between 1968 and 1979, examining 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 on a variety of outcomes, including skyjackings and non-skyjacking incidents, such as hostage taking, barricades, and attacks on diplomats
Enders, W., Sandler, T., & Cauley, J. (1990). UN Conventions, Terrorism and Retaliation in the Fight Against Terrorism: An Econometric Evaluation. <i>Terrorism and Political Violence</i> , 2(1), 83	Evaluated multiple interventions and outcomes used an interrupted time series approach for events between 1968 and 1988, examining 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, W., & Sandler, T. (1993). The Effectiveness of Antiterrorism Policies: A Vector-Autoregression-Intervention Analysis. <i>The American Political Science Review</i> , 87(4), 829-844	Used interrupted time series analysis (improving upon techniques used in previous works) to examine 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
Brophy-Baermann, B., & Conybeare, J.A.C. (1994). Retaliating Against Terrorism: Rational Expectations and the Optimality of Rules Versus Discretion. <i>American Journal of Political Science</i> , 38(1) (Feb), 196-210	Used 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, W., & Sandler, T. (2000). Is Transnational Terrorism Becoming More Threatening? <i>Journal of Conflict Resolution</i> , 44, 307-332	Expanded their analysis to terrorism events between 1970 and 1996. Studied 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 on outcomes which measure the type of person-based destruction – death events, wounded, and non-casualties per quarter
Barros, C. (2003). An Intervention Analysis of Terrorism: The Spanish Eta Case. <i>Defence and Peace Economics</i> , 14(6), 401-412	Used intervention analysis specifically on information regarding assassinations and kidnappings conducted by the Spanish terrorist group ETA between 1968 and 2000 collected by Abadie and Gardeazabal (2001). Barros studied the effects of different political ideologies in power, police and military expenditures as well as increases in foreign investment on these incidents using a vector autoregression time series framework

<i>Table 3</i> Summary of findings			
Intervention category	«Works»	«No statistically discernible effect»	«Was Harmful»
	<i>Terrorism declined</i>	<i>Levels of terrorism did not change</i>	<i>Terrorism increased</i>
Metal Detectors and Security Screening	12	3	13
Fortifying Embassies and Protecting Diplomats	3	19	6
Increasing the Severity and Certainty of Punishment	0	2	0
U.N. Resolutions and Conventions Against Terrorism	2	2	0
Military Strikes	0	5	6

Metal detectors and security screening

Interventions which increased detection and hardened targets of potential terrorism through airport security have been one 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. The success of metal detectors in airports has often been widely recognized as effective in reducing hijackings and twelve findings support this. However, thirteen findings indicate

that metal detectors showed a harmful effect – that is, that after the intervention there was an increase in terrorism events. A closer examination indicated that the difference between beneficial and harmful findings were the outcomes measured; for findings that were beneficial, all outcomes measured were hijacking events while the findings that showed metal detectors increased terrorism measured non-hijacking offenses. As Cauley and Im (1988) and Enders and Sandler (1990; 1993) have pointed out, this difference may point to substitution or displacement effects of airport security on other types of terrorism.

Fortifying embassies and protecting diplomats

In terms of interventions which physically fortified embassies as well as increased security and protection for diplomats, many findings were non-significant and close to 0, indicating that there is little scientific evidence of these interventions being effective in reducing terrorism against such targets. It is interesting to note that fortifying embassies and protecting diplomats through increased security at embassies aren't as effective as metal detectors in airports, even though both were situational crime prevention measures intended to harden targets. This could be the result of airports being more secure and stable environments than the outsides of embassies housed in other countries or diplomats on the move. Thus, perhaps a lesson from both of these findings is that target hardening strategies may be much more effective in more controlled environments.

Increasing the severity of punishment

Landes (1978) provided the only findings concerning increasing the severity of punishment for hijackers who were apprehended. 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.

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 are, in theory, supposed to provide a general deterrent effect on terrorism by establishing international norms which affect or strengthen national policies against terrorism. However, only Enders et al. (1990) discovered that a United Nations resolution against aerial hijackings (that also supported the use of metal detectors in airports) appeared effective in reducing the number of hijacking events in both the short and long term. However, 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.

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 & 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 in our review indicated that the attack on Libya resulted in a statistically significant *increase* in the number of terrorist attacks in the short run. However, 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, both the retaliations in Libya as well as Israeli attacks on the PLO seem to increase attacks on the United States, United Kingdom, and Israel generally.

Changes in political governance

Finally, we examined findings that could be grouped in the general intervention category of «changes in 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. Their findings indicate an uncertainty about whether the existence of harsher parties on terrorism as well as the end of the Cold War may increase terrorism events. 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.

Discussion

Both our preliminary review and the more in-depth Campbell systematic study presented important lessons in terms of the evidence base of counter-terrorism policy. Most importantly, commonly used interventions, 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 worse, have been shown to be harmful and increase the likelihood of certain types of terrorism to occur. 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-run, less lethal threats and other activities that do not use substantial resources.

Perhaps what is equally (if not more) interesting is what we *didn't* find from our review. Most interventions have never been evaluated, which speaks to the lack of an evidence base for counter-terrorism policy. For example, we found articles discussing such measures as anti-terrorism products for personal use, arrest and imprisonment in Guantánamo Bay, assassinations, blast resistant luggage, detection devices for biological or chemical weapons, diplomatic and/or third party efforts, education, emergency response preparedness, the use of gas masks for biological agents, hostage negotiation, laws against terrorism,

emergency preparedness for bioterrorist attack, prosecution strategies, psychological counseling, religious interventions, seal/tamper proof devices, other social and economic responses, and use of the media, to name only some. Many of these interventions (and more) have become part of our daily lives and discourse, but we have no idea whether or not they fulfill the promise of reducing terrorism, terrorism-related risk or harm as there is no research to support these interventions. Indeed, some of these interventions may also have collateral and unintended effects of reducing civil rights or our quality of life. It is often asserted that perhaps people are willing to give up some of their rights for safety and security if programs are effective. The problem is that we are unsure of whether these interventions for which people are giving up their rights or dignities actually work.

Thus, the findings strongly confirm our initial speculation about the state of counter-terrorism strategies. There has been a proliferation of counter-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 or continue to use programs that we know are ineffective or harmful. Counter-terrorism strategies are clearly not evidence-based. The most important policy recommendation to emerge from this review is that counter-terrorism policies need to be evaluated for their effectiveness or at least be better informed by existing scientific evaluations.

Conclusions

These findings lead us to the key focus of this paper: recommendations for improving the evidence base of counter-terrorism policy. Specifically, we focus our recommendations on three types of decision-makers: government agencies and agents attempting to counter terrorism, government agencies funding counter-terrorism efforts, and researchers and policy-makers involved in the evaluation of these interventions. Each recommendation is intended to improve the evidence base of counter-terrorism research and also more generally suggests improvements in the infrastructure of counter-terrorism policy and research.

To government agencies and policy makers generating and implementing counter-terrorism measures

Counter-terrorism policy needs to be rational, effective, and cause as little harm as necessary. There is only one way to determine whether counter-terrorism strategies are effective – by conducting methodologically valid evaluations of those strategies. It is clear that current counter-terrorism policies, strategies and tactics lack this evidence base. In other words, programs are being used without any knowledge, understanding, or even attempts to determine whether they are effective. Government agencies and policy makers should be aware that they are appropriating large amounts of monies to programs which do not reduce terrorism, in some cases increase terrorism, or have no effect at all. Thus, government agents need to pay attention to scientific research about counter-terrorism programs when making policy choices or in the least, encouraging, facilitating, and conducting scientific evaluations of these programs. 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 determine the effectiveness of a particular strategy.

This need for evaluation research requires that both science and scientists are welcomed into governmental counter-terrorism enterprises to facilitate evaluations. Indeed, researchers understand that much information associated with terrorism is classified and requires security protections. 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 such fears. Through an ongoing program of assessment and review, in areas such as crime prevention, police tactics, and drug abatement programs, the value and limitations of many police interventions have now been documented. This has led to major improvements in establishing programs that can improve police effectiveness. For counter-terrorism efforts, government agencies should consider extending clearances to evaluation researchers so they can study the effectiveness of policies and assist in more effective and efficient government spending.

To government agencies funding counter-terrorism research and development

Currently, there has been an increase in funding for counter-terrorism research and development, much of which has been directed towards the creation of new programs, technologies, strategies and tactics to counter terrorism. The problem is that we already have a number of counter-terrorism programs that are not evidence-based. Funding is needed for the evaluation of existing counter-terrorism programs, rather than the creation of new programs. As was stressed in the previous *Psicothema* issue of 2006 (vol. 18, n° 3) dedicated to crime prevention, and as Garrido, Farrington & Welsh (2008) point out, program proliferation in an environment which lacks an evidence base is not only bad policy but also could be dangerous, as indicated in the findings above.

Additionally, «process descriptions» which describe a program's procedure or determine if procedures were carried out according to a plan are not outcome evaluations. Funded evaluations should therefore include outcome evaluations which are 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, funding agencies that set stronger methodological requirements for evaluation 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 both outcome evaluations as well as the use of higher-quality evaluation designs.

Along the same lines and, especially with regard to terrorism, funding agencies should encourage the discovery of alternative or improved ways to gauge effectiveness of counter-terrorism programs. For example, can experimental and other types of quasi-experimental designs (in addition to time-series) be used to evaluate some programs (perhaps those used to respond and manage terrorism and terrorism-related byproducts)? Or, are there other methods that might be useful in evaluating the effects of programs on rare events? Such methodological explorations are fundamental to developing an evidence base for counter-terrorism strategies given the nature of the subject matter.

To researchers studying terrorism

Only a small fraction of terrorism research appears to be empirical in nature and an even smaller proportion are evaluations, both of which researchers need to focus more on. Yet, increasing the amount of empirical evaluation research is a major enterprise which requires that evaluation researchers work with policy makers and practitioners to improve the research infrastructure for counter-terrorism evaluations. This includes pressing for access to terrorism-related data that goes beyond incident based information and includes the assessment of intervention strategies. This type of work can only be pursued when there are partnerships between researchers, policy makers and practitioners that set as their goal the interpretation and dissemination of findings. Additionally, researchers need to explore ways in which terrorism strategies can

be analyzed and different types of methodologies that may be useful.

Evaluation research can serve as a moderating and rational effect on rash policy responses based on moral panic and fear. This is important, as rash and unscientific policies can lead to other social negatives, including the violation of personal and human rights as well as individual humiliation, both of which could potentially lead to more terrorism. Scientists and policy-makers can moderate the proliferation of bad policies with more research that evaluates both the outcome effectiveness as well as the social, political, economic, or psychological effects of these interventions. The call for a larger evidence base for counter-terrorism policy is not a criticism of policy makers; it is a joint responsibility between evaluation researchers and policy makers to facilitate and create.

References

- Abadie, A., & 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.
- Boruch Robert, F., Snyder, B., & DeMoya, D. (2000). The importance of randomized field trials. *Crime and Delinquency*, 46, 156-80.
- Brophy-Baermann, B., & 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., & Im, E. (1988). Intervention policy analysis of skyjackings and other terrorist incidents. *The American Economic Review*, 78(2), 27-31.
- 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.
- Cullen, F.T., & 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., & Smith, P.C. (2000). *What works: Evidence-based policy and practice in public services*. London: Policy Press.
- Enders, W., & Sandler, T. (1993). The effectiveness of antiterrorism policies: A vector-autoregression-intervention analysis. *The American Political Science Review*, 87(4), 829-844.
- Enders, W., & Sandler, T. (2000). Is transnational terrorism becoming more threatening? *Journal of Conflict Resolution*, 44, 307-332.
- Enders, W., Sandler, T., & Cauley, J. (1990). UN conventions, terrorism and retaliation in the fight against terrorism: An econometric evaluation. *Terrorism and Political Violence*, 2(1), 83.
- Garrido, V., Farrington, D.P., & Welsh, B. (2008). Crime prevention: More evidence-based analysis. *Psicothema*, 20(1), 1-3.
- Guinnessy, J., & 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.
- 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.
- Landes, W.M. (1978). An economic study of U.S. aircraft hijackings, 1961-1976. *Journal of Law and Economics* 21, 1-31.
- Lum, C., & Yang, S. (2005). Why do evaluation researchers in crime and justice choose non-experimental methods? *Experimental Criminology*, 1(2), 191-213.
- Lum, C., Kennedy, L., & Sherley, A. (2006[a]). *The effectiveness of counter-terrorism strategies: A Campbell Systematic Review*. The Campbell Collaboration. Available online at: http://www.campbellcollaboration.org/doc-pdf/Lum_Terrorism_Review.pdf.
- Lum, C., Kennedy, L., & Sherley, A. (2006[b]). Are counter-terrorism strategies effective?: The results of the Campbell Systematic Review on counter-terrorism evaluation research. *Journal of Experimental Criminology*, 2, 489-516.
- 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.
- Miller, R. (1988). The literature of terrorism. *Terrorism*, 11, 63-87.
- Nutley, S., & Davies, H.O. (1999). The fall and rise of evidence in criminal justice. *Public Money and Management*, 19, 47.
- Prunckun, H., & Mohr, P. (1997). Military deterrence of international terrorism: An evaluation of Operation El Dorado Canyon. *Studies in Conflict and Terrorism*, 20, 267-280.
- Romano, T. (1984). *Terrorism: An analysis of the literature*. Dissertation, Fordham University, Department of Sociology, Criminology and Penology.
- Schmid, A.P., & Jongman, A.J. (1988). *Political terrorism: A new guide to actors, authors, concepts, databases, theories and literature*. Amsterdam: North-Holland Publishing Company.
- 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., & MacKenzie, D.L. (Eds.) (2002). *Evidence based crime prevention*. London: Routledge.
- Silke, A. (ed.). (2004). *Research on terrorism: Trends, achievements and failures*. New York: Taylor and Francis.
- Silke, A. (2005). Fire of Iolau: 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.
- Weisburd, D., Lum, C., & 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., & 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.
- Welsh, B., & Farrington, D. (2002). What works, what doesn't, what's promising and future directions. In Sherman, L., Farrington, D., Welsh, B., & Mackenzie, D. (Eds.): *Evidence-based crime prevention*. London: Routledge.
- Welsh, B., & Farrington, D. (2006). Closed-Circuit Television Surveillance. In Welsh, B. & Farrington, D.: *Preventing crime: What works for children, offenders, victims and places*. Dordrecht, Netherlands: Springer.