The effectiveness of surveillance technology

What intelligence officials are saying

Cayford, Michelle; Pieters, Wolter

DOI
10.1080/01972243.2017.1414721

Publication date
2018

Document Version
Final published version

Published in
Information Society

THE INFORMATION SOCIETY
2018, VOL. 34, NO. 2, 88–103
https://doi.org/10.1080/01972243.2017.1414721

Routledge
Taylor & Francis Group

The effectiveness of surveillance technology: What intelligence officials are saying

Michelle Cayford and Wolter Pieters

Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

ABSTRACT

In recent years, Western governments have come under sharp criticism for their use of surveillance technology. They have been accused of sweeping up massive amounts of information without evidence of the technologies being effective in improving security. The view of critics is clear, but what do intelligence officials themselves say? This paper analyzes statements of intelligence officials in the U.S. and U.K. from 2006 to 2016, examines what criteria officials use in their discourse on effectiveness, and investigates how considerations of cost and proportionality factor into the equation. It identifies seven measures of effectiveness in the statements of intelligence officials, and finds that cost, though rarely discussed, is the driver behind *formalized* evaluations of surveillance programs.

ARTICLE HISTORY
Received 5 October 2016
Accepted 27 November 2017

Introduction

Surveillance technology is pervasive in our society today, leading to fierce debate between proponents and opponents. Government surveillance, in particular, has been brought increasingly under public scrutiny, with proponents arguing that it increases security, and opponents decrying its invasion of privacy. Since the Snowden leaks, critics have loudly accused governments of employing surveillance technologies that sweep up massive amounts of information, intruding on the privacy of millions, but with little to no evidence of success. And yet, evaluating whether surveillance technology increases security is a difficult task. How does one measure the value of one bit of intelligence that contributes to the greater whole? How do we measure the role of intelligence in informing decision-makers?

This paper focuses on what intelligence officials in the U.S. and U.K. themselves say about the effectiveness of surveillance technology. In their own words, what are the criteria for evaluating whether a particular piece of surveillance technology meets the goal that motivated its deployment? Even in the absence of explicit evaluations, intelligence bodies must constantly make judgments about effectiveness to determine if they will continue to use and redeploy a particular surveillance technology. This evaluation of effectiveness may be implicit, but it is there.

This study does not examine the veracity of officials' statements, nor does it determine whether or not a particular surveillance technology is actually effective. Our approach is not to question the truth of what officials say, nor to judge actual effectiveness, but to delve into the meaning and significance of intelligence officials' statements, to identify values they place on effectiveness and the measures they use to assess it, and their reasoning. This study is not hostile to security forces, but an attempt to honestly understand what considerations intelligence officials take into account when they speak about the effectiveness of surveillance technology. Because so much surrounding surveillance technology is controversial, how it is discussed matters.

The paper proceeds as follows: after briefly addressing related work, terminology is defined and the research methods of this study are described. Thereafter what intelligence officials are saying about effectiveness of surveillance technology is examined. Next, statements about cost are analyzed to understand how officials factor it into their assessments is followed by an analysis of what they are saying with regard to proportionality. Lastly, officials' statements regarding effectiveness, cost, and proportionality are considered together and critiqued, and recommendations offered for a more holistic approach.

Related Work

This section covers the parts within the broad literature on security and surveillance literature that are most relevant to this paper. One of these is the privacy and security debate, which has been at the forefront in recent

CONTACT Michelle Cayford m.r.cayford@tudelft.nl Safety and Security Science, TU Delft, P.O. Box 5015, Delft, NL 2600GA.

years. Here, the issue of effectiveness is central, as the basis for arguing for the use of any given surveillance program must be that it is effective in increasing security. Moreover, any proportionality judgment must consider the privacy intrusion of the program against its effectiveness. This literature can be classified into three categories: 1) actual effectiveness 2) belief of effectiveness and 3) statements of effectiveness.

Many authors have written on the privacy concerns raised by the intersection of government surveillance and modern technology (e.g. Greenwald 2014, Berghel 2013, Morgan 2014, Monahan 2016, Bigo et al. 2013). Some of these authors accuse Western democracies of exaggerating the threat of terrorism to justify mass surveillance (e.g. Greenwald 2014) and also of exaggerating its role in preventing terrorist activity (e.g. Bergen et al., 2014).

A study of the U.S. Department of Homeland Security's fusion centers, which are meant to facilitate information sharing among relevant government agencies, finds that there is confusion regarding legal and accountability frameworks that should be followed, resulting in mission creep and privacy violations (Regan & Monahan, 2013). Another study on these fusion centers finds not only that mission creep occurs and thereby potential privacy violations, but also that the centers are ineffective in their primary tasks of information awareness and sharing (Monahan & Palmer 2009). Ineffectiveness at achieving initial goals leads to subsequent efforts to make the centers useful and effective in other ways, and thereby mission creep occurs with changes in centers' tasks.

The largest number of studies on effectiveness examine the actual effectiveness of surveillance technology – assessments and measurements of whether or not a given security program accomplishes its security goal. There is a significant body of work on evaluation of the effectiveness of counterterrorism measures. Lum et al. (2007), van Dongen (2009), and van Um and Pisoiu (2011) identify the numerous challenges of performing an effectiveness evaluation, propose approaches to measuring effectiveness, and underline the lack of research in this field. Drakos and Giannakopoulos (2009) establish a formal statistical framework to determine the probability of authorities stopping a terrorist incident over time and the probability of human and property loss. Predictive data mining has been analyzed as a counterterrorism method and argued to be ineffective (Jonas & Harper, 2006). One study purports to analyze the effectiveness of counterterrorism approaches in six countries, but in reality is an historical account of the terrorism in each country and the counterterrorism policies and practices put in place by the government (Alexander, 2006). A second study by van Dongen (2015) constructs a new framework for evaluating counterterrorism policies and examines whether there is a relation between the type of terrorist organization and the effectiveness of the counterterrorism approaches applied to combating it.

There are some government-related reports that address the question of the actual effectiveness of security measures. They include two in-depth reports by the U.S. Privacy and Civil Liberties Oversight Board (PCLOB), which discuss measures of effectiveness used by intelligence officials and present the Board's own conclusions about the effectiveness of NSA surveillance programs (PCLOB, Jan. 2014 and July 2014). Another report by the Congressional Research Service makes some rather obvious points, such as increased expenditures in counterterrorism does not necessarily result in progress (Perl, 2007).

One way to evaluate the actual effectiveness of surveillance technology is a cost-benefit analysis. Mueller and Stewart (2011) use risk assessment to determine the risk of a terrorist attack and then gauge whether the costs of security measures are outweighed by the benefit of a likely-prevented attack. They argue that the enormous amounts of spending in the U.S. on anti-terrorist measures far outweigh the benefits gained. In the law enforcement realm Edwards et al. (2014) point to a lack of evaluation of online data mining technology. Hewitt (2014) looks at the actual effectiveness of law enforcement tactics for dealing with the various types of terrorism in America and concludes that whether or not different types of tactics are effective depends on the type of terrorism. Additionally, Ekblom (2010), develops a framework for crime prevention and security in the community and Sproles (1999), a method for establishing measures of effectiveness that can be applied to any field.

A small body of work deals with the actual effectiveness of specific kinds of surveillance technology. There are two RAND Corporation reports on measuring effectiveness in specific contexts. The first one is on assessing the effectiveness of U.S. Air Force remotely piloted aircraft, more commonly known as drones (Lingel et al., 2012). The second one is on measuring the effectiveness of border security (Willis et al., 2010). Tsvetovat and Carley (2006) evaluate several information gathering programs to determine their effectiveness in mapping the connections between members of covert organizations. Stewart and Mueller (2011) conduct a cost-benefit analysis of a specific piece of surveillance technology – full body scanners.

Lastly, there is literature on CCTV cameras and their effect on crime. As Gill and Spriggs (2005) point out, studies on the effectiveness of CCTV have arrived at

various conclusions. Some find that CCTV has some effect in reducing crime (Armitage et al. 1999; Farrington et al. 2007), while others find it has no effect at all (Ditton and Short 1999; Gill et al. 2006; Phillips 1999). Certain recent studies, however, have identified conditions in which CCTV operates most effectively (Gill and Spriggs 2005, Caplan et al., 2011). In two systematic reviews using meta-analysis Welsh and Farrington (2003, 2009) found that CCTV is most effective at reducing crime in car parks. Similarly, Ratcliffe (2006) notes that CCTV literature, as well as CCTV studies referenced in his paper, point to CCTV working best in small, defined spaces and against property crimes rather than violent crimes.

All the above-mentioned studies deal with actual effectiveness. Effectiveness can also be studied from the standpoint of belief. Bruce Schneier argues that terrorist attacks are actually rare, and that security theater – implementing security measures that are not proven to actually increase security, but give the public a sense of security – plays into the hands of the terrorists by treating them as legitimate military opponents and overreacting to their fear tactics (Schneier, 2009). Putz (2012) argues the exact opposite: what is important is that security practices appear to be effective, not necessarily that they are. If the public, including terrorists, does not know which practices are actually effective but perceives them to be effective, terrorist actions may be deterred, which will result in actual effectiveness.

In the U.K., a report produced by the National Policing Improvement Agency discusses research on community policing, which encourages the public to cooperate with the police and be socially responsible, leading to crime reduction in a cost-effective manner. According to Myhill and Quinton (2011), public trust in the police rests less on the perception of the police effectively fighting crime, than on the belief that the police acts fairly when dealing with the public. Trust contributes to overall effectiveness in the long term.

The last category of studies concerns statements of effectiveness – what people operating in the security domain say regarding effectiveness. How do they discuss the effectiveness of their work? How do they define success? Sanders et al. (2015) argue that intelligence-led policing in Canada has cultivated a culture wherein police services define their success in terms of accountability rather than outcomes. In a study on financial surveillance Amicelle (2011) notes that the "results count less than demonstrations of codes of conduct" (p.173). Coaffee and Fussey (2015) examine the resilience, security, and surveillance discourse, and conclude that since 9/11 there has been a shift in vocabulary from "security" to the more positive "resilience," while the fundamental focus of security has remained the same.

Our current study focuses on statements by intelligence practitioners regarding the effectiveness of their surveillance technology. It fills a gap by focusing specifically on intelligence agencies and on their surveillance technology. It assesses their statements in order to understand the measures by which practitioners evaluate effectiveness, as well as their justification for their surveillance programs.

Terminology

In all the literature reviewed for this study only two papers defined "effective." The EU SURVEILLE project considers a surveillance technology to be effective when it "has the technical capacity to deliver the intended security goals, and when employed for a defined goal within the necessary context achieves the intended outcome" (van Gulijk et al., 2013, p.3). Van Dongen (2015) defines effective as "an impact that is desirable in the eyes of the state… and can be observed in the terrorist actor" (p.85). We take "effective" to be an impact that is desirable and can be observed as contributing towards the sought-after security goal. Since this paper focuses on the realm of intelligence, which deals with gathering information, effectiveness comes to mean obtaining sought-after information, which then contributes to achieving the overall desired security goal. Here it is important to note that *effectiveness* – whether or not a surveillance technology achieves its security goal – differs from *performance* – the technology's technical capacity and ability to function correctly (e.g. Currie and Stiefvater, May 2003).

"Surveillance technology," as used in this paper, can in principle, include a range of technologies, from wiretaps to drones to satellites to all manner of cameras (hidden, CCTV, etc.). In the material analyzed, however, intelligence officials are primarily referring to systems dealing with communications data (surveillance systems monitoring phone calls, emails, Internet activity, etc.). These are the types of systems Snowden exposed.

Intelligence practitioners, as evidenced in this study, use the term "surveillance programs" or "collection programs" when speaking about the systems that perform surveillance. "Surveillance technology," as such, is not spoken of. The term "program" is broadly used, referring to either one kind of surveillance technology or to multiple technologies used together to collect a particular type of data, or data from a particular source. For example, traffic may be intercepted from the Internet, filters applied to this data, and certain data selected out and stored for a given period of time. All of this would be

referred to as a "program" (Omand, Mar. 2015). Intelligence practitioners also talk of "tools" – individual components that make up a program. Because intelligence officials do not make clear distinctions between these terms, in this paper, the words "program" and "technology" and "tools" are used interchangeably, unless otherwise specified.

Mentions of "agency" refer to intelligence bodies, such as the NSA, CIA, and GCHQ. Intelligence "agency" and intelligence "body," therefore, are used synonymously.

Methodology

Non-classified statements of intelligence practitioners were analyzed to address our research questions: How do intelligence practitioners articulate effectiveness? How are factors of cost and proportionality taken into account in their discussion?

The question of effectiveness is not determined in a vacuum. There are inevitably other factors at play. Even if a technology is determined to be effective, the ultimate decision of whether or not to use it is also based on considerations such as expense and proportionality. Cost, although not discussed at length in the material analyzed, was confirmed in interviews to be a factor that affects the choice of surveillance technology. Proportionality was heavily discussed by officials in the materials analyzed. This paper considers effectiveness in a strict sense (whether or not the technology achieves the sought-after security goal), as well as with cost and proportionality in an overall effectiveness evaluation.

This study analyzed statements made from 2006 to 2016 by directors and former directors of the U.S.'s National Security Agency (NSA) and Central Intelligence Agency (CIA), and the U.K.'s Government Communications Headquarters (GCHQ), in the form of speeches, congressional and parliamentary testimonies, articles, and books. The 2006–2016 timespan was chosen to have a good amount of time (7 years) prior to the Snowden leaks to enable a substantive comparison between the directors' pre and post Snowden statements. 2006 also marks the beginning of General Hayden's term as director of the CIA.

Since 2006 there have been two directors of the NSA (Alexander and Rogers), four of the CIA (Hayden, Panetta, Patraeus, and Brennan), and three of GCHQ (Pepper, Lobban, and Hannigan). Of note are two former directors who have produced a number of documents and made statements during this time frame – Sir David Omand, director of GCHQ from 1996–1997, and General Michael Hayden director of the NSA from 1999–2005 and director of the CIA from 2006–2009. Their status as former directors is particularly interesting as it gives them a bit more liberty to speak about the work of their respective agencies. The reader will note that much of the U.K. analysis is based on Omand documents and statements, as he has been a much more prolific writer and speaker on intelligence issues than his GCHQ counterparts.

The criterion for selecting speeches, statements, articles, etc. was that the subject matter included elements of effectiveness, cost, or proportionality. This determination was made by looking at the titles and scanning the material. For example, an article by Omand on developing national resilience was not chosen, while an article on ethical guidelines for using intelligence for public security was selected. Likewise, there are many testimonies given by NSA and CIA directors before congressional committees on current threats the U.S. faces. These were not selected. Selected material was then searched for the following keywords: effective, efficient, success, works, surveillance, technology, cost, budget, finance, privacy, proportionate, and balance.

The following web pages were searched for relevant material: the NSA and CIA statements, speeches, and testimonies pages, the GCHQ speeches page, the transcripts and public evidence of the Intelligence and Security Committee of the U.K. Parliament page, the subcommittees (NSA, cybersecurity, and CIA) of the House of Representatives Permanent Select Committee on Intelligence pages. Additionally, a Google scholar and Scopus search was performed for each director by name, and statements by other practitioners (non-directors of the GCHQ, NSA, and CIA and those from other agencies) were extracted from a public workshop transcript and two reports issued by the PCLOB on NSA programs.

In addition, the following officials were interviewed: a former senior U.S. government official (Interview 5, 2016), a former U.S. intelligence officer (Interview 3, 2015), a former senior police officer from a U.K. counter-terrorism network (Interview 1, 2015), Her Majesty's Inspector of Constabulary (Interview 6, 2015), and a former senior investigation officer of the U.K. North West Counter Terrorism Unit (Interviews 7 & 8, 2015). All interviewees were involved in the output of intelligence. Those from the U.K. were from the police force, so while they worked on terrorism cases, sometimes alongside MI5 agents, they did not speak directly from the perspective of an intelligence agency and thus they are referenced less in the paper. Two additional interviewees from an international context – a high-level advisor and recipient of intelligence in Estonia (Interview 2, 2015) and a former cryptanalyst with the Dutch Military Intelligence and Security Service (Interview 4, 2015) – provided additional insight into the world of intelligence. The

responses of the interviewees supported, in private conversations, what other intelligence practitioners were saying in public settings.

In total, 42 documents were analyzed and 8 interviews were conducted. The gathered materials were then categorized to identify the ways of assessing effectiveness and discussions associated with effectiveness, cost, and proportionality. Any differences in the treatment of effectiveness between the U.S. and the U.K. were also analyzed.

What Intelligence Officials are Saying About…

What Intelligence is

Typically, law enforcement conducts investigations after a crime occurs, while intelligence bodies collect information in advance of and even independently of any "event" occurring. Today, particularly with the advent of modern terrorism, these lines are blurred and overlap occurs in reality. Law enforcement now also engages in performing intelligence work to stop terrorist attacks and dismantle terrorist cells before an event occurs. One fundamental difference that does remain is that intelligence agencies inform policy. That is, "intelligence exists solely to support policy makers in myriad ways" (Lowenthal, 2012, p. 2) by providing them "deep-reached, nuanced, strategic appreciations" (Council on Foreign Relations, 2015).

Directors of the GCHQ, NSA, and CIA have underlined that the basic purpose of intelligence is to help decision-making, while making clear that this information does not dictate what action should be taken. It is for government officials to decide what action to take based on the intelligence they are provided (Pepper, Dec. 2010). As noted by Omand (2014, p.14), "The most basic purpose of intelligence… [is] to help improve the quality of decision-making by reducing ignorance… to help improve the quality of decisions, not to guarantee that result." In other words, intelligence work engages in gathering information to inform. It is not in the business of investigating suspicious people, as is the case in law enforcement. General Hayden states very frankly and clearly: "'Suspicionless surveillance' doesn't make sense. I'm not a law enforcement officer. I don't suspect anybody; I'm collecting information to keep the country safe. NSA doesn't just listen to 'bad' people; it listens to interesting people. The information is what we're pursuing" (W&L Symposium, 2015, YouTube video).

As the purpose of intelligence is to inform decision-making, it operates within the realm of politics. It cannot be separated from politics and politicians. "The policy maker is not a passive recipient of intelligence but actively influences all aspects of intelligence" (Lowenthal 2012, p.2). Effectiveness therefore must be considered in this context with its particular complexities. The intelligence system is set up to be neutral – intelligence agencies deliver the intelligence and policy makers decide what, if anything, to do with it. However, policy makers can cherry-pick intelligence – select the intelligence that suits their political agenda and ignore the rest (Interview 5, 2016).

Strategic vs. tactical intelligence

There are two kinds of intelligence – strategic and tactical (or operational) – which are quite different from one another. Tactical operations are more specific – targeted at specific individuals or groups. Here individuals are put under surveillance because they are, for example, suspected of plotting a terrorist attack or of being spies passing classified information to foreign countries. There is a defined beginning and end to this type of surveillance. Intelligence collected on a strategic level, however, is more broad – conducted against a foreign government or military for an unspecified time period, for example. While the goal is to gain information on the target entity's activity, what will be discovered is unknown. Strategic intelligence "determine[s] the nature of the threat," while tactical intelligence "relates to a specific operation" (Her Majesty's Inspectorate of the Constabulary 2015, p. 27).

Since the objectives of strategic and tactical intelligence are different, the criteria for assessing their effectiveness must also be different. When it comes to tactical intelligence, officials care if the job gets done. In other words, if surveillance technology aids in providing information that shows that someone was passing along classified information, or that a certain individual belonged to a terrorist group, it would be considered effective (Interview 5, 2016).

When it comes to strategic intelligence, however, whose basic purpose is to inform, it is much harder to evaluate effectiveness. One former intelligence official suggested that this is probably not even possible. The goal of strategic intelligence is information dominance. The goal is in "as many ways as possible, to gather as much information as possible, to solve as many problems as possible" (Interview 5, 2016). Further complicating the issue, gathering strategic intelligence involves big systems, such as satellites. These systems are tasked with multiple intelligence-gathering goals at the same time, and their missions may change over time. Thus, any evaluation would involve measurement against multiple goals (Interview 5, 2016). (One could imagine a large system for gathering Internet communications having the same challenge.)

Intelligence generated through a piece of surveillance technology may prove to be key years after it was

obtained. Intelligence officials often state that intelligence is about putting together the pieces of the puzzle. So, by itself a piece of intelligence may not seem so significant, but put together with multiple other pieces of intelligence it could become crucial.

Effectiveness

One of the findings of this study was how much officials had to say on the respective topics of effectiveness, cost, and proportionality. There was a lot of material on proportionality, while very little about cost. The varying lengths of the following sections – effectiveness, cost, proportionality – is reflective of this finding.

Effectiveness as a topic in and of itself was rarely discussed. It was addressed in a general way, such as a particular technique being characterized as an "effective program." Only in one instance did intelligence officials specifically treat the question of effectiveness in relation to certain surveillance programs. In other instances, surveillance techniques were discussed in terms of proportionality and privacy, with officials arguing the usefulness of the program.

The author's own communication with a former senior U.S. government official indicates that there is some consideration of effectiveness within the intelligence community. However, the evaluation may be more of the performance of the intelligence agencies' themselves rather than of specific surveillance technologies. This is in keeping with the larger observation that officials tend to speak of evaluation of an intelligence agency as a whole and not of specific surveillance technology or programs used by it (Hayden, 2008). Generally, officials tend to see intelligence to be a product of aggregate work, as it typically requires contributions from multiple sources and also good analysis. As some interviewees noted, surveillance technology may collect golden information, but if it is not properly analyzed it means nothing (Interview 2, 2015; Interview 5, 2016). So, officials shy away from an evaluation of the technology itself, preferring to assess the final analyzed product. Therefore, to evaluate effectiveness one must evaluate what the agency produces, which means an evaluation of the agency as a whole.

Although officials rarely addressed effectiveness directly, seven measures of effectiveness could nonetheless be identified from the data analyzed. These measures can be grouped into three categories: counting, documents/cases, and organizations.

Counting

1. *Thwarted attacks*

In the wake of the Snowden leaks NSA director General Alexander cited the number of terrorist activities – 54 – that had been disrupted as a result of information collected by surveillance programs operating under Section 215 of the Patriot Act and Section 702 of the Foreign Intelligence Surveillance Act of 1978 (FISA). This was meant to show that the programs were effective – success stories showcased as evidence that these surveillance programs were keeping the country safe.

However, since General Alexander's statement, American intelligence officials have emphasized that counting success stories is not a measure with which one should evaluate a program's effectiveness. Robert Litt, General Counsel, Office of the Director of National Intelligence, cautioned: "[I]ndividual success stories are not the way to evaluate a collection program and its utility" (PCLOB, July 2014). Similarly, Rajesh De, General Counsel, National Security Agency, argued: "I think the absolute wrong question is how many plots did this tool stop" (PCLOB, July 2014). U.K. officials share this view. According to Omand, "It is less a question of how many terrorist attacks, criminal plots and cyber attacks have been stopped because of specific interception of terrorist intent in their communications and much more the unique contribution digital intelligence sources make to the intelligence jigsaw and the painstaking process of 'discovery' of terrorist cells and involved individuals" (Omand, Mar. 2015, p.5).

2. *Lives saved*

The number of lives saved from detecting terrorist communications and subsequently thwarting terrorist plots is another measure of effectiveness. For instance, Lobban et al., (2013) note that U.K. intelligence agencies depend on "the fantastic work that GCHQ do to detect terrorist communications. That leads to us finding terrorist plots that we would not otherwise find, that we are then able to thwart, which leads to lives being saved" (p.17).

3. *Terrorist (and criminal) organizations destroyed*

In writing about targeted killings using drones, Hayden argues that it has been a very effective program. Drones are outfitted with cameras, both still and video, making them a form of surveillance equipment. They can also be equipped with missiles, as in the case of this targeted killing program. Hayden states that the program has not only disrupted terrorist plots, but it also "reduced the original Qaeda organization along the Afghanistan-Pakistan border to a shell of its former self" (Hayden, 2016, online). Greatly reducing or destroying such an organization is judged to be an effective outcome.

Documents/cases

4. *Output*

Intelligence officials explicitly cite output that an intelligence agency generates for policy makers as a

measure of effectiveness. In the U.S. the Office of the Director of National Intelligence (ODNI) has done studies to determine if resources are effectively allocated within the intelligence community. One metric used is reports generated – by quantity and by quality. With regard to quantity, ODNI looks at the number of reports generated that cover the collection priorities[1] that the intelligence community has been given. If a surveillance technology has been the collection source of numerous reports it is judged to be effective. As regards quality, the ODNI looks at the nature of the information and "its utility towards a whole variety of national priorities." That is, it identifies the collection source of important intelligence reports (Mr. Litt, quoted in PCLOB July 2014, p.65).

In a discussion on equipment interference (the installation of malware on a device to activate the microphone or camera, collect location information, etc.) Omand stated that such capability is of "inestimable value to the intelligence agencies… Some 20% of GCHQ's output benefits from that kind of technique" (Omand, Dec. 2015, pp.603–604). Omand clearly considers equipment interference to be an effective technique, and the measure of effectiveness he uses is how much this technique has contributed to the agency's overall output. Output here is not defined, but it can be presumed to mean GCHQ's output to its customers, such as intelligence reports. Also of note is the percentage figure – contribution to 20% of output is considered to be an acceptable number to qualify to be effective.

Another such effectiveness figure was given by General Hayden in a reference to the NSA's phone metadata program. Following the Snowden leaks on this program, President Obama changed the provision for contact chaining[2] from 3 to 2 hops out from the original number. Hayden gave his professional judgment that this change "preserved about 85% of the effectiveness of the program. And in the real world where politics matter, that's OK" (W&L Symposium, 2015, YouTube video). Although he judges the program to be effective, he does not give the metric by which he makes this determination. Changing the number of hops lessens its effectiveness, but it is still 85% effective, which is a good number according to Hayden.

a. Use in criminal cases

The U.K. Home Secretary stated that communications data was used in 95% of criminal cases. Similarly, Omand characterizes communications data as "a very important investigative tool" in terrorist trials. Here the use of data in criminal cases is seen as an indication of the effectiveness of the surveillance program that gathered the data (Omand, Oct. 2014, p.6).

Organizations

5. *Context*

Intelligence professionals also look at how surveillance programs complement other tools. NSA's Rajesh De points out that "all intelligence tools are used in complementary fashion with one another and to isolate one particular tool and evaluate its effectiveness in isolation probably doesn't do us justice as to what's valuable and what's not" (PCLOB July 2014, p.65).

6. *Support*

Practitioners judge the support rendered to other agencies via intelligence collected by a particular surveillance system to be a measure of effectiveness. Omand argues that bulk collection is effective. One indication that this is so is that other intelligence agencies and law enforcement depend on GCHQ and its bulk interception. The largest part of the U.K. intelligence budget goes to GCHQ, so if these other agencies did not find GCHQ's activities, including bulk interception to be useful, they would argue that they could use the money better than GCHQ (Omand, Oct. 2014).

7. *Informed policy maker*

One interviewee stated that ultimately effectiveness is decided by policy makers' needs, not by intelligence (Interview 5, 2016). Another interviewee stated that intelligence is positively evaluated if the customer (i.e. policy maker) feels informed (Interview 2, 2015). In the same vein, a GCHQ director speaks of assessments based on "the quality of service" (Pepper, Jan. 2010, p.94).

Analysis of effectiveness

In this section general observations on the evaluation of effectiveness in the intelligence community are analyzed, as well as the seven measures of effectiveness that were drawn from the data. The latter are analyzed according to themes identified.

General and complex

There was a tendency to speak of evaluations of an intelligence agency as a whole and not of specific surveillance technology or programs used by it. This reflects intelligence officials' belief that surveillance programs should be evaluated in their complementary relationship to one another, and not in isolation. If the effectiveness of a group of intelligence tools is being assessed as a unit, where does that group end? At what point is a surveillance technology excluded, deemed to not be complementary to other technologies within the group? It is easy to see how eventually we arrive at evaluating the agency as a whole, with the use of all the surveillance technologies it is equipped with. This tendency of evaluating intelligence bodies as a whole means that

surveillance technology itself is not formally assessed for its effectiveness.

Further, our findings underlined the fact that intelligence operates within the realm of politics. It cannot be separated from politics and politicians. Ultimately, then, effectiveness is not tied to whether the surveillance program is delivering the security goal, but to whether the policy makers are getting the intelligence they need to feel informed. This may be judged according to how the information fits into their political agenda, which can be influenced by the image they wish to project to their electorate.

To count or not to count

The findings show an interesting tension between counting and not counting. Intelligence officials state that evaluation of intelligence should not be based on counting of successful cases. Yet, there is still a tendency to be drawn toward metrics that involve counting – number of thwarted attacks, number of lives saved, number of terrorist organizations destroyed, and number of criminal cases that drew on intelligence gathered. This is indicative of at least two things. One, the inclination to measure effectiveness quantitatively – an easy way to establish the success of something is to count off the number of times it has given positive results. Two, it points to the difference between strategic and tactical intelligence. Strategically there are no success stories. There is information that informs a country's leaders about a particular government, economic situation, strife, military programs, etc. But it does not translate into a success story that can be counted. Tactically, however, it can be possible to quantify how many terrorists have been identified, how many plots thwarted, etc. It seems intelligence practitioners tailor effectiveness evaluations to different kinds of intelligence gathering.

Measures and manner

The measure of *context* points to the way in which practitioners believe evaluations should be conducted. Considering the context and complementary manner in which surveillance technologies are used is a necessary condition for assessing effectiveness.

The data on GCHQ's measure of *support* actually highlights two separate metrics of effectiveness. One, the support GCHQ renders other agencies by giving them valuable information gathered via its surveillance systems. Two, the other agencies' support of GCHQ and its surveillance systems as evidenced by their acceptance of the allocation of the largest portion of the U.K.'s intelligence budget to it.

In the other points professionals make about effectiveness, some could be considered to overlap. For example, *lives saved* carries us back to thwarting attacks and the issue of counting. If counting attacks thwarted is not a measure of effectiveness, then *lives saved* cannot be either because this remains in the realm of measuring effectiveness by counting success stories. If counting disrupted plots *is* used as a measure then this raises the question of whether measuring lives saved is not double counting, counting both the attack prevented and the lives saved from the attack, both of which would be the result of the same intelligence.

Effectiveness as a percentage

The percentages used by officials when speaking about effectiveness are interesting in that they point to possible acceptable thresholds of effectiveness. If a surveillance program contributes to 20% of an agency's overall output, this is considered a good figure, presumably well over the threshold. Likewise, a program that maintains 85% of its effectiveness is considered acceptable. In a similar vein, communications data are said to be used in 95% of British criminal cases, which is seen to be an indicator of success, pointing to the effectiveness of the underlying programs. In other words, while presumably a technology is effective in at least some percentage of the cases, in order to be considered effective it must reach a persuasive threshold of percentage of successful cases.

Effectiveness mapped

The measures for the effectiveness of surveillance technology mentioned by intelligence officials are mapped in Figure 1. Measures that involve counting are separated by a dotted line. The darker the dotted line, the more clearly the measure falls into the "counting" category. If the evaluation is related to strategic intelligence, where "counting" measures are less relevant, the figure could be cut along any one of these dotted lines to eliminate 1, 2, or all types of "counting" measures. *Support* is bi-directional as indicated by the arrow, showing the support a surveillance technology provides to an outside agency, as well as that outside agency's support of that technology. Evaluations should be done in consideration of technologies being complementary to one another, as shown by the "context" domain.

In summary, we could say that formalized evaluations of effectiveness seem to be performed for intelligence agencies as a whole and not of specific surveillance technologies. Measures of effectiveness identified in intelligence officials' statements, however, indicate that *output* is a central measure and that effectiveness is established when the percentage of output to which that program contributes, reaches a persuasive threshold, which remains undefined. Intelligence officials appear to value counts of successful cases as a measure of effectiveness

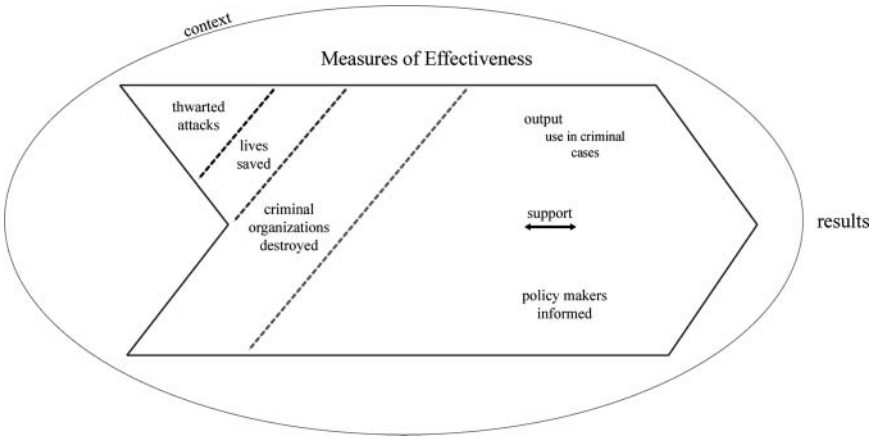

Figure 1. Evaluating effectiveness of surveillance technology as described by intelligence officials.

for tactical intelligence, but not for strategic intelligence. On the other hand, effectiveness is not necessarily seen as being tied to the security goal but instead to satisfying the information needs of the policy makers. This, arguably, could negate the need for any other measures.

Cost

In the documents and speeches analyzed the subject of cost did not appear that often. The few times that it did appear it was in general statements, such as, "Our Congressional overseers… work every day to ensure that American taxpayer dollars are being spent effectively and efficiently to keep our country strong" (Brennan, 2014, non-paginated transcript). In the one venue in which effectiveness was specifically treated as a subject, James Baker, General Counsel, Federal Bureau of Investigation (FBI) said: "we have an obligation… to spend our time and spend our money on programs that are effective and not be wasting our time on things that are not" (quoted in PCLOB July 2014, p.85).

It is, however, no secret that intelligence agencies are constrained by budgets (even if that budget is large). In interviews conducted by the author, all officials emphatically affirmed that cost is most certainly a factor when it comes to deploying surveillance technologies. This refers not only to the actual money spent, but also to manpower. In reference to tactical operations in the U.K., "resources are so limited and the volume of potential terrorists is so high that your threshold has got to be very high to put surveillance" (Interview 1, 2015). Abroad, agencies want to ensure that a program or operation is giving them intelligence, or is leading to supplying intelligence. If not, it will be shut down (Interview 3, 2015). A manager wants to know that his resources are being applied effectively. But how do you know if the money is spent effectively? How do you know what is the right number of people to put on a job? (One interviewee posed these rhetorical questions, without answering them.) To some degree cost can determine the quality of information obtained. That is, a more powerful satellite will give clearer photos with a higher pixel number, giving a very clear image of what is on the ground (Interview 5, 2016).

Among these general statements, two revealed more specifics. An NSA surveillance program, Trailblazer, was eventually shut down due to being far over budget. General Hayden, NSA director at the time, told a Senate committee that the costs, "were greater than anticipated, to the tune of, I would say, hundreds of millions" (Mayer, May 2011, online). In the U.K., Omand noted that "the content of an encrypted message does not represent a cost-effective target for the authorities" (Omand, Mar. 2015, p.3). The cost of attempting to read the content of an encrypted message is too high relative to any information that might be gained.

Analysis of cost

On the one hand, it is not rocket science that cost, an important factor, is considered in determining and evaluating surveillance programs. On the other hand, it is rocket science – or at least secret science – how this determination is made. Besides affirming that intelligence agencies are limited by resources and manpower and therefore that their analysts and agents focus only on the top tier targets, this study did not yield much insight into the fine grain, nitty-gritty of how intelligence officials consider cost in the overall assessment of effectiveness of surveillance programs.

One insight it did yield is that, at times, the judgment regarding whether a program was effective or not was implicitly connected to its cost. For example, ODNI performs evaluations of surveillance programs to check whether resources are allocated effectively. Here effectiveness is not being considered in a vacuum but in the context of cost.

Proportionality

Human rights lawyers and the media heavily stress the question of proportionality; the crux of the criticism levied against the NSA and the GCHQ, in particular, centers on this point. The accusations are that these agencies employ surveillance technologies that gather huge amounts of data, but the number of cases in which this data has been shown to protect the public is minuscule, and even here evidence of effectiveness is questionable. Consequently, proportionality looms large in the statements of intelligence officials as they respond to these accusations.

Addressed by law

Time and again, when intelligence officials are asked questions related to proportionality and privacy, they refer to the law. For them, the law and oversight of intelligence bodies establish what is proportional. They then act within these parameters. Therefore, they themselves do not need to make judgments about whether a surveillance program is proportional.

Directors of these intelligence bodies view their job as providing intelligence, while working within the legal framework. What the law itself says is an issue for politicians to debate, and if they so choose, to change. Lobban, former director of GCHQ, states, "[Legislation] is an issue for politicians and not for us. We are not law makers. There are strict criteria in the law which provides safeguards to protect privacy to the maximum extent possible… If Parliament chooses to have a debate, fine by me. If Parliament chooses to change the laws, so be it" (Lobban et al., 2013, p.18). In the U.S., Hayden reflects this same view when he says that the space within which the CIA operates "is defined by the policymakers that we all elect and by the laws our representatives pass" (Hayden, 2007, non-paginated transcript).

… and by human beings

And yet, although the law establishes boundaries for intelligence agencies, officials recognize that within these boundaries there is a human element that determines proportionality. Firstly, there is the person signing the request – Home Secretary, Foreign Secretary, or judge in the U.K., or the Foreign Intelligence Surveillance Court (commonly known as FISA Court) in the U.S. Secondly, directors themselves can make judgments of proportionality. Hayden, as NSA director on 9/11 did this very thing. He says that prior to 9/11 certain communications were not considered valuable but thereafter they were deemed critical to national security. In other words, what he viewed as reasonable (or proportionate) on the morning of Sept. 10 was different than what he saw as reasonable on the afternoon of Sept 11. For instance, after 9/11, collection of American phone metadata was determined to be lawful and proportionate (W&L Symposium, 2015; Hayden, 2006).

Thus, while the law establishes proportionality (or establishes that any surveillance must be proportional) on one level, within that legal framework personal judgments are made on what is, in fact, proportionate. The aftermath of 9/11 is a good example of how the judgment of where this line falls can change.

The issue of mass surveillance

Intelligence officials are firm in their stance that what they do is not mass surveillance. "Mass surveillance is about pervasive observation or monitoring of the entire population or a substantial sector of it. Observation implies observers, human beings who are examining the thoughts and actions of the population" (Omand, Mar. 2014, p.3).

1. *The amount of data collected*

With the issue of proportionality comes the question of how much data intelligence agencies collect, particularly in regards to the collection of communications data off the Internet. On the one hand, modern digital communications have generated massive flows of information. Hayden argues that the only way for agencies to handle these volumes of data is to perform bulk collection. (Hayden, May 2014) On the other hand, even in collecting this data in bulk, the NSA itself states that it touches a mere 1.6% of Internet traffic. Of that 1.6%, only 0.025% is selected for review and seen by an analyst. In effect, NSA analysts see only 0.00004% of the world's Internet traffic (NSA, 2013). In a similar vein, Omand strongly denies the accusation that the GCHQ is processing data about everybody (Omand, Dec. 2015).

The argument here is that while the NSA and GCHQ collect a significant amount of data, it is a small fraction of the world's Internet traffic, and it is not everyone's data. This is in contrast to mass surveillance, which would begin with collecting everyone's data.

2. *What they collect*

What intelligence agencies collect, as governed by law, further shows, according to officials, that they are not conducting mass surveillance and that this surveillance is proportionate.

Both the U.S. and Britain have very strict laws about the collection of their own citizens' data. But there are some significant differences between them in the rules governing the collection of foreigners' data. In the U.S. there is a strong distinction between "U.S. persons" – American citizens and foreigners who are in the U.S. – and non-U.S. persons. Because of this distinct difference between these two categories of people, what foreign

intelligence collects related to these groups differs greatly. U.S. persons are protected by the Constitution, which provides protection against unreasonable searches and seizures. Accordingly, collection of U.S. persons' data by the CIA and NSA is not allowed. (The FBI is the agency that investigates people within the U.S. suspected of criminal activity.) Anybody else is fair game. As Hayden put in stark terms, "Your privacy is simply not the concern of the NSA director" (Hayden, Feb. 2014, online).

The U.K., however, does not make such a distinction in its collection of data. The legislation governing the interception of communications by law enforcement is also applied to all the intelligence agencies. If someone poses a security threat, British intelligence will seek to intercept that person's communications. If someone is not a security threat and is not in contact with someone who is, intelligence agencies are not permitted to intercept their communications. According to Lobban, "We are not entitled to. That is true, actually, whether you are British, if you are foreign, and wherever you are in the world" (Lobban et al., 2013, p.15).

U.K. law – specifically RIPA 2000 – dictates what is classified as content of a communication and what is not, and therefore what can and cannot be seen without a warrant. A GCHQ analyst is authorized to look at the IP address of the suspect computer, the user's email address, when and where the communication originated, and the server identity being accessed. Therefore, they can see that the user accessed Google, but not what they searched for. In Internet communications data, everything beyond the first slash (e.g. beyond www.google.com/) is considered content and a warrant from the Secretary of State must be obtained to access it (Omand, Mar. 2014).

3. *How they collect*

British intelligence officials have provided some detail as to how they collect data. Several GCHQ directors have made the analogy of the Internet as an enormous haystack and of GCHQ looking for needles inside that haystack. GCHQ tries to gather hay from the parts of this haystack it has access to which could potentially hold needles or parts of needles. Queries are then designed to draw the needles out of this part of the haystack. The surrounding hay may have been intercepted, but it will not be looked at. Only that for which there is a specific authorization is looked at (Lobban et al., 2013, p.13).

Data collection performed by these agencies is "not indiscriminate collection of data willy-nilly" (Omand, Dec. 2015). Omand repeatedly draws a distinction between what is collected by computers based on algorithms created to search for certain communications, and what is, from that collection, selected according to certain criteria (laid out in search warrants) and then seen by an analyst. Computers search through the bulk data to find the sought-for communications. When they find it, the data is pulled out. This filtered data is what is kept and what the analyst sees. Such selection of data is based on what a warrant has authorized and this guarantees that privacy is respected. According to Omand, it is an "unwarranted assumption that access in bulk to large volumes of digital communications (the 'haystack') in order to find the communications of intelligence targets (the wanted 'needles') is evidence of mass surveillance of the population, which it is not" (Omand, Mar. 2015, pp.8–9). Buffering, or keeping the bulk data for a day or two while the computers search it, is necessary because it is not technologically possible to do a real-time analysis of all the bulk data (Omand, Oct. 2014). Omand maintains that it is a "highly discriminating, selective use" of surveillance tools in order to find the communications data of suspects and that bulk access "is not being used as some giant fishing expedition" (Omand, Oct. 2014, pp.2 & 5).

The U.S. government takes this same view that the temporary acquisition of data in order to search it according to specific "selectors"[3] does not constitute mass surveillance (Presidential Policy Directive, 2014). Furthermore, the government is not able to access or make use of the collected communications other than to determine if they contain a selector (PCLOB, July 2014).

Analysis of proportionality

American and British officials rely on slightly different arguments to make their case to their respective publics. The U.S. officials, targeting American audiences, outline the strict rules governing the collection of U.S. persons' data. This is to reassure the U.S. public that the NSA and CIA are not looking at their data – they are not allowed to and oversight mechanisms ensure that they do not. To the rest of the world the message is that while our laws do not restrict our gathering of your data, our limited resources do. We are going after potentially dangerous targets and do not have the time or desire to waste energy on lower tier targets.

In the U.K., since the same laws govern the collection of citizens' and foreigners' data, the issue becomes the necessity of bulk collection – the haystack. A responsible intelligence agency should engage in bulk collection, as there is no other technical way to find potential terrorists and other security threats. The individual's privacy is ensured because computers do automated searching; human eyes only see what is selected. The distinction made here between collection and selection raises the question of whether or not the collection of data is surveillance. The argument seems to be that since human eyes are not looking at it, the data in question is not under surveillance. Therefore, proportionality is not an issue. This also brings into question whether or not this

program needs to be effective. If collection by computers is not surveillance, then the equipment performing the collection is not surveillance technology and therefore does not need to be evaluated for effectiveness.

The so-called balance

In assessments of the effectiveness of a surveillance technology, what do intelligence officials say regarding the balance between the cost, proportionality, and actual effectiveness of the technology?

In 2002[4] General Hayden addressed this issue in a Congressional testimony. He spoke of finding the right balance between protecting security and protecting liberty, and asked Committee members to talk to their constituencies and "find out where the American people want that line between security and liberty to be" (Hayden, 2002, non-paginated transcript). Where this line is drawn has far-reaching consequences for how the NSA carries out its mission – the focus of its activities, the standard for conducting surveillance, the type of data it can collect, how it collects the data, the rules for retaining and disseminating U.S. persons' information (Hayden, 2002).

Three years later, in 2005, he wrote an article addressing this very topic and entitled, "Balancing Security and Liberty." Here Hayden calls this a "pressing" question. He goes on to addresses this question in the context of the NSA sharing data while at the same time protecting U.S. privacy rights. Hayden says: "The oversight structure… has ensured that the imperatives of national security are consistent with democratic values" (Hayden, 2005, p.251). In other words, the law and oversight ensure that the right balance is struck.

British intelligence officials, on the other hand, are not in favor of the term "balance." They point out that it is not a choice between security and privacy, but that the two go together – in order to enjoy privacy, citizens must firstly have a secure society. Security "provides the fundamental basis upon which other rights can be more easily secured. A State that is suffering insecurity will be badly placed to deliver the protection of other rights, including privacy" (Omand, Feb. 2014, p.1). Here the notion of a balance between the two ultimately is problematic because it implies that having more of one automatically means less of the other. GCHQ directors emphasize that they believe that their job is to provide both – deliver security while protecting privacy (Lobban, 2014).

It is the combination of "practices, procedures, laws and regulations" that "helps to ensure that intelligence activity is legal, ethical and effective" (Omand et al., 2012, pp.18–19). The delivery of security while maintaining proportionality is achieved through laws and regulations in concert with security practices and procedures. Periodic review of all of the above contributes to maintaining this balance (Omand, Feb. 2014).

Intelligence officials recognize that public perception plays a role in overall effectiveness. That is, in a democracy, if the public does not trust their nation's intelligence agencies due to the employment of certain surveillance programs, ultimately the operation of the intelligence agency will be greatly hindered (Omand, 2005). The government needs "to lead in education for the public, because this will affect the overall effectiveness of the security strategy" (Runciman, 2012, p.37).

Analysis of 'balance'

How this so-called balance is struck or how this dual-mission of attaining effectiveness and privacy is achieved takes us back to the discussion on proportionality. On both sides of the Atlantic intelligence officials agree that law and oversight ensure that this so-called balance is kept. Where they differ is the notion of balance. In the U.S. the balance between effectiveness and privacy is talked about freely, while in the U.K. it is rejected because it implies that the furtherance of one is at the expense of the other. British officials seek to deliver both effectiveness and privacy simultaneously.

The terms officials use when addressing the so-called balance – "security and liberty" or "security and privacy" – give reason for pause. The problem here is that the discussion essentially remains in the realm of proportionality. What is needed is a true discussion of balance (or triple-mission) that gives adequate attention to all three elements – effectiveness, cost, and proportionality. In other words, the three elements should be treated in triple tandem. Rather than just speaking of providing "security," the debate should be sharpened to discuss the effectiveness of the surveillance technology in achieving the security goal. Firstly, assess if the technology is effective. If it is effective in achieving the given security goal, then ask, is it proportional? Additionally, the question of

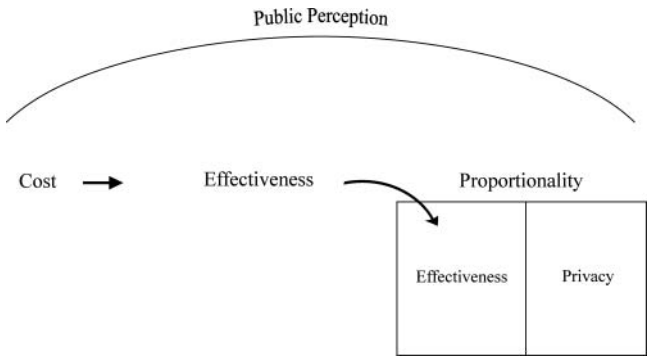

Figure 2. Components of overall effectiveness.

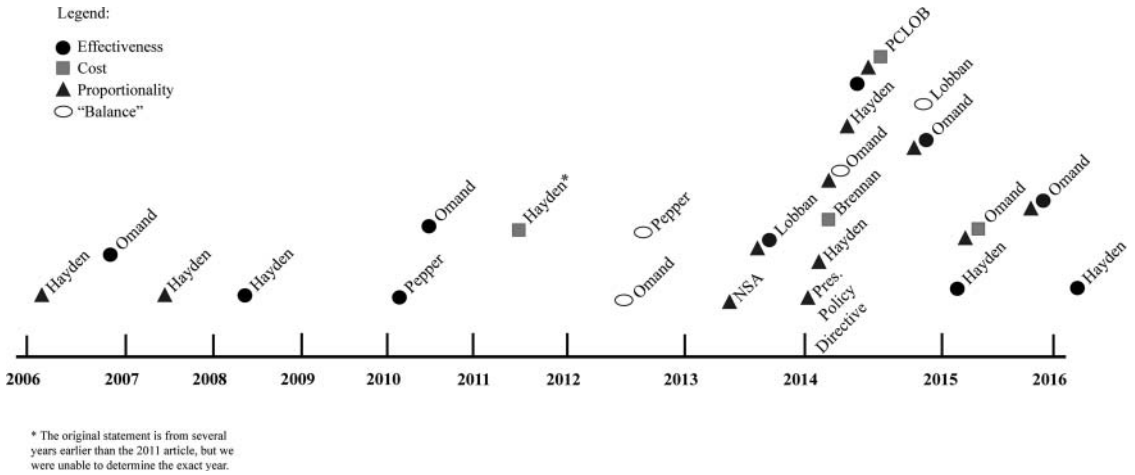

Figure 3. Officials' statements related to effectiveness, cost, proportionality, and "balance."

cost should be considered in this equation. Is the budget expended justifiable for the security goal obtained? This is the kind of triple tandem in which all three elements must be taken into account.

If the case of the ODNI can be taken as our prototype (being the most concrete example we have of performed evaluations of surveillance programs), then cost is the ultimate driver of formal evaluations of surveillance systems. What drives governments to evaluate their surveillance technologies is not a desire to assess effectiveness, but to determine if funds are being appropriately spent. This reality is expressed in Figure 2. Here the role cost considerations play in prompting effectiveness concerns is indicated by the arrow. Effectiveness both stands on its own and feeds into proportionality, which encompasses both effectiveness and privacy. In other words, surveillance technology is evaluated for its effectiveness in advancing security. This effectiveness then becomes part of proportionality in determining what is appropriate in terms of using the effective surveillance technology and simultaneously protecting citizens' privacy. Public perception is shown to span all three categories of effectiveness, cost, and proportionality, as how the public perceives how much is spent on surveillance technology, how effective that technology is, and whether its use is proportional, all ultimately influence the overall effectiveness of surveillance programs.

Conclusion

This paper analyzes U.S. and U.K. intelligence officials' statements in the 2006 – 2016 time period regarding the effectiveness of surveillance technology – including statements on cost and proportionality, which play into determinations of overall effectiveness. Figure 3 plots over time intelligence officials' statements related to effectiveness, cost, proportionality, and "balance."[5]

The key points of intelligence officials' statements on the effectiveness of surveillance technology are that, it is extremely difficult, if not impossible to evaluate the effectiveness of surveillance programs. Intelligence work is like putting together pieces of a puzzle – multiple seemingly insignificant parts come together to form an important and critical picture. When it comes to effectiveness, it becomes difficult to evaluate one small piece of the puzzle that by itself seems insignificant but is necessary for the completion of the picture. Further, the purpose of intelligence is to inform policy makers, to improve the quality of their decision making. Measuring the impact of strategic intelligence on the decision-making process it informs is difficult.

Seven measures of effectiveness were drawn from intelligence officials' statements: thwarted attacks, lives saved, criminal organizations destroyed, output, context, support, and informed policy-maker. Officials argue that counts of successful cases should not be a measure of effectiveness. Yet, the tendency to do just that shows up in the value they attribute to surveillance programs. This indicates a difference in evaluation of tactical vs. strategic intelligence. Counting successful cases seems to have some merit with officials as a measure of effectiveness of surveillance technology employed for tactical intelligence purposes, but not for strategic intelligence. With all the measures, the percentage of instances that serves as the threshold for deeming a technology to be effective becomes important, as presumably any technology will be effective in at least some cases.

Officials state that the law determines the boundaries of proportionality, and oversight mechanisms ensure that the intelligence bodies stay within these limits. Further, there is the distinction between bulk data collected by computers, and limited selected data seen by human eyes. Lastly, cost considerations drive governments to perform formalized evaluations of surveillance programs.

Addressing the empirical question of how intelligence officials articulate effectiveness is a necessary starting point for any subsequent dialogue regarding the use of surveillance technology. Other stakeholders are also making statements about the effectiveness of surveillance technology. Privacy advocates are, in particular, critical of intelligence agencies' use of surveillance programs and their interpretation/application of proportionality. To arrive at a consensus on how such technology should be used and regulated, oversight bodies, the public, privacy advocates, and intelligence agencies must begin with an understanding of how the others value and measure effectiveness. This study provides such an understanding with regard to intelligence officials. Further studies could investigate how the other groups address and treat effectiveness.

Notes

1. Collection priorities are the policy issues and areas that policy makers task intelligence bodies to collect intelligence on (Lowenthal 2012, p.57).
2. Contact chaining refers to the phone numbers in contact with the number under investigation. If phone number 333 is being looked at, all those who were in contact with this number will also be looked at. That is one hop. The second hop entails studying all the numbers that were in contact with the numbers under the first hop.
3. "A selector must be a specific communications facility that is assessed to be used by the target, such as the target's e-mail address or telephone number" (PCLOB, July 2014, p.32)
4. No material was found within the time frame of 2006-2016 in which American officials addressed the "balance" question. These two documents from 2002 and 2005 were therefore included in the analysis. One document by Omand from 2005 was also included.
5. Statements made by interviewees are not included in this figure because the discussions were prompted by the author.

Acknowledgements

The authors wish to acknowledge Pieter van Gelder for his contribution of ideas to this paper.

References

Alexander, Y., ed. 2006. *Counterterrorism strategies: Successes and failures of six nations*. 1st ed. Washington, DC: Potomac Books, 2006.

Amicelle, A. April 2011. Towards a 'new' political anatomy of financial surveillance. *Security Dialogue* 42(2)(April 2011):161–78.

Armitage, R., G. Smyth, & K. Pease. 1999. Burnley CCTV evaluation. In *Surveillance of public space: CCTV, street lighting and crime prevention*, ed. K. Painter and N. Tilley, 225–250. Monsey, NY: Criminal Justice Press.

Bergen, P., D. Sterman, E. Schneider, and B. Cahall. January 2014. *Do NSA's Bulk Surveillance Programs Stop Terrorists?* New America Foundation.

Berghel, H. July 2013. Through the PRISM Darkly. *Computer* 46(7):86–90.

Bigo, D., S. Carrera, N. Hernanz, J. Jeandesboz, J. Parken, F. Ragazzi, and A. Scherrer. 2013. National Programmes for Mass Surveillance of Personal Data in EU Member States and Their Compatibility with EU Law. *European Parliament*.

Brennan, J. 2014. Remarks at the council on foreign relations. March 11.

Caplan, J. M., L. W. Kennedy, and G. Petrossian. September 2011. Police-monitored CCTV cameras in Newark, NJ: A quasi-experimental test of crime deterrence. *Journal of Experimental Criminology* 7(3):255–74.

Coaffee, J., and P. Fussey. (February 2015). Constructing resilience through security and surveillance: The politics, practices and tensions of security-driven resilience. Edited by Myriam Dunn Cavelty, Mareile Kaufmann, and Kristian Søby Kristensen. *Security Dialogue* 46 (1):86–105.

Council on Foreign Relations. 2015. *Homeland Security Implications of ISIS Attacks*. http://www.cfr.org/homeland-security/homeland-security-implications-isis-attacks/p37320.

Currie, N., and K. Stiefvater. May 2003. Counter-terrorism technology assessment and methodology study. Final Technical Report. Rome: Air Force Research Laboratory Information Directorate.

Ditton, J., and E. Short. 1999. Yes, It Works, No, It Doesn't: Comparing the Effects of Open CCTV in Two Adjacent Scottish Town Centres, In *Surveillance of public space: CCTV, street lighting and crime prevention*, eds. K. Painter and N. Tilley, 201–224. Monsey, NY: Criminal Justice Press.

Drakos, K., and N. Giannakopoulos. April 2009. An econometric analysis of counterterrorism effectiveness: The impact on life and property losses. *Public Choice* 139(1–2):135–51.

Edwards, M., A. Rashid, and P. Rayson. June 2014. A systematic survey of online data mining technology intended for law enforcement. *ACM Computing Surveys* 48(1).

Ekblom, P. 2010. *Crime prevention, security and community safety using the 5Is Framework*. Houndmills, Basingstoke; New York, NY: Palgrave Macmillan.

Farrington, D. P., M. Gill, S. J. Waples, and J. Argomaniz. 2007. The effects of closed-circuit television on crime: Meta-analysis of an English national quasi-experimental multi-site evaluation. *Journal of Experimental Criminology* 3:21–38.

Gill, M., and A. Spriggs. February 2005. *Assessing the impact of CCTV*. Home Office Research, Development and Statistics Directorate.

Gill, M., A. Rose, K. Collins, and M. Hemming. 2006. Redeployable CCTV and drug-related crime: A case of implementation failure. *Drugs: Education, Prevention and Policy*, 113(5):451–460.

Greenwald, G. 2014. *No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State*. First Edition. New York, NY: Metropolitan Books/Henry Holt.

Hayden, M. 2005. Balancing security and liberty: The challenge of sharing foreign signals intelligence. *Notre Dame Journal of Law, Ethics and Public Policy* 19(1):247–60.

Hayden, M. February 2014. Beyond Snowden: An NSA reality check. *World Affairs Journal*. http://www.worldaffairsjournal.org/article/beyond-snowden-nsa-reality-check.

Hayden, M. May 2014. *Munk Debates – State Surveillance – Pre-Debate Interview*. https://www.munkdebates.com/debates/state-surveillance.

Hayden, M. 2008. Remarks at the Atlantic Council. November 13.

Hayden, M. 2007. Remarks at the Council on Foreign Relations. September 7.

Hayden, M. 2002. *Statement for the Record before the Joint Inquiry of the Senate Select Committee on Intelligence and the House Permanent Select Committee on Intelligence*.

Hayden, M. 2006. *What American Intelligence & Especially the NSA Have Been Doing to Defend the Nation*. National Press Club, January 23.

Hayden, M. V. 2016. *To keep America safe, embrace drone warfare*. The New York Times, February 19, http://www.nytimes.com/2016/02/21/opinion/sunday/drone-warfare-precise-effective-imperfect.html.

Her Majesty's Inspectorate of the Constabulary 2015. An Inspection of the National Crime Agency. Available at: https://www.justiceinspectorates.gov.uk/hmicfrs/wp-content/uploads/an-inspection-of-the-national-crime-agency.pdf (accessed November 30, 2017).

Hewitt, C. January 2014. Law enforcement tactics and their effectiveness in dealing with American Terrorism: Organizations, autonomous cells, and lone wolves. *Terrorism and Political Violence* 26(1):58–68.

Interview 1, Former senior police officer from UK counter-terrorism network. April 20, 2015.

Interview 2, High level recipient of intelligence in Estonian government. June 10, 2015.

Interview 3, Former Intelligence Officer of U.S. Intelligence Community. July 28, 2015.

Interview 4, Former cryptanalyst with Dutch Militaire Inlichtingen- en Veiligheidsdienst. August 27, 2015.

Interview 5, Former senior U.S. government official. January 19, 2016.

Interview 6, Otter, S. Her Majesty's Inspector of Constabulary, Interview, May 7, 2015.

Interview 7, Smith, M. Detective Superintendent, Greater Manchester Police (former Senior Investigation Officer for Operation Pathway, UK North West Counter Terrorism Unit), Interview, May 28, 2015.

Interview 8, Smith, M. Detective Superintendent, Greater Manchester Police (former Senior Investigation Officer for Operation Pathway, UK North West Counter Terrorism Unit), Interview, September 9, 2015.

Jonas, J., and J. Harper. 2006. Effective counterterrorism and the limited role of predictive data mining. *Policy Analysis*. Washington D.C.: Cato Institute, December 11.

Lingel, S. L., L. Menthe, B. Alkire, K. Henry, S. A. Grossman, R. A. Guffey, and E. Wu. 2012. *Methodologies for analyzing remotely piloted aircraft in future roles and missions*. Documented Briefing. Santa Monica, CA: RAND.

Lobban, S. I.. "Valedictory Speech," October 21, 2014. https://www.gchq.gov.uk/speech/sir-iain-lobbans-valedictory-speech-delivered.

Lobban, S. I., A. Parker, and S. J. Sawers. Nov. 7, 2013. *Intelligence and Security Committee of Parliament: Uncorrected Transcript of Evidence*.

Lowenthal, M. M. *Intelligence: From secrets to policy*. 5th ed. Los Angeles: SAGE/CQ Press, 2012.

Lum, C., L. W. Kennedy, and A. Sherley. January 2007. Are counter-terrorism strategies effective? The results of the Campbell systematic review on counter-terrorism evaluation research. *Journal of Experimental Criminology* 2 (4):489–516.

Mayer, J. May 23, 2011. *The secret sharer*. The New Yorker. www.newyorker.com/magazine/2011/05/23/the-secret-sharer.

Monahan, T. August 7, 2016. Built to Lie: Investigating technologies of deception, surveillance, and control. *The Information Society* 32(4):229–40.

Monahan, T., and N. A. Palmer. December 2009. The emerging politics of DHS fusion centers. *Security Dialogue* 40 (6):617–36.

Morgan, S. A. 2014. Security vs. liberty: How to measure privacy costs in domestic surveillance programs. Master's thesis, Naval Postgraduate School.

Mueller, J. E., and M. G. Stewart. 2011. *Terror, security, and money: Balancing the risks, benefits, and costs of homeland security*. Oxford , New York: Oxford University Press.

Myhill, A., and P. Quinton. 2011. *It's a fair cop? Police legitimacy, public cooperation, and crime reduction*. National Policing Improvement Agency.

National Security Agency. August 9, 2013. The National Security Agency: Missions, Authorities, Oversight and Partnerships.

Omand, D. December 2005. Countering International Terrorism: The Use of Strategy. *Survival* 47(4):107–16.

Omand, D. March 2015. *Understanding digital intelligence and the norms that might govern it*. Global Commission on Internet Governance. Chatham House. https://www.cigionline.org/publications/understanding-digital-intelligence-and-norms-might-govern-it.

Omand, D. February 2014. Evidence for the intelligence and security committee of parliament.

Omand, D. October 2014. *Intelligence and Security Committee of Parliament: Privacy and Security Inquiry, Public Evidence Session 8*.

Omand, D. December 2015. *Joint Committee on the Draft Investigatory Powers Bill – Oral Evidence*.

Omand, D. 2014. The future of intelligence: What are the threats, the challenges and the opportunities? In *The future of intelligence: Challenges in the 21st century*. Oxon: Routledge.

Omand, D., J. Bartlett, and C. Miller. 2012. A balance between security and privacy online must be struck…. *Demos*.

Pepper, S. D. January 1, 2010. The business of Sigint: The role of modern management in the transformation of GCHQ. *Public Policy and Administration* 25(1):85–97.

Pepper, S. D. December 2010. *Testimony on Iraq*.

Perl, R. March 12, 2007. Combating terrorism: The challenge of measuring effectiveness. *CRS Report for Congress*. Congressional Research Institute.

Phillips, C. 1999. A review of CCTV evaluations: Crime reduction effects and attitudes towards its use. In *Surveillance of public space: CCTV, street lighting and crime prevention*, ed. K. Painter and N. Tilley, 123–155. Monsey, NY: Criminal Justice Press.

"Presidential Policy Directive – Signals Intelligence Activities." *Whitehouse.gov*, January 17, 2014. https://www.whitehouse.

gov/the-press-office/2014/01/17/presidential-policy-directive-signals-intelligence-activities.

Privacy and Civil Liberties Oversight Board (PCLOB). July 2, 2014. Report on the Surveillance Program Operated Pursuant to Section 702 of the Foreign Intelligence Surveillance Act.

Privacy and Civil Liberties Oversight Board (PCLOB). January 23, 2014. Report on the Telephone Records Program Conducted under Section 215 of the USA PATRIOT Act and on the Operations of the Foreign Intelligence Surveillance Court.

Putz, O. April 1, 2012. From non-places to non-events: The airport security checkpoint. *Journal of Contemporary Ethnography* 41(2):154–88.

Regan, P. M., and T. Monahan. 2013. Beyond counterterrorism: Data sharing, privacy, and organizational histories of DHS fusion centers. *International Journal of E-Politics* 4(3) (33):1–14.

Runciman, B. June 1, 2012. A Bigger Haystack…. *ITNOW* 54 (2):36–37.

Sanders, C. B., C. Weston, and N. Schott. July 2015. Police innovations, 'Secret Squirrels' and accountability: Empirically studying intelligence-led policing in Canada. *British Journal of Criminology* 55(4):711–29.

Schneier, B. November 2009. Beyond Security Theater. *Schneier on Security*. https://www.schneier.com/blog/archives/2009/11/beyond_security.html.

Sproles, N. 1999. *Measures of effectiveness: The standards for success*. PhD Thesis, University of South Australia.

Stewart, M. G., and J. Mueller. January 16, 2011. Cost-benefit analysis of advanced imaging technology full body scanners for airline passenger security screening. *Journal of Homeland Security and Emergency Management* 8(1).

Tsvetovat, M., and K. M. Carley. November 8, 2006. On effectiveness of wiretap programs in mapping social networks. *Computational and Mathematical Organization Theory* 13 (1):63–87.

van Dongen, T. December 2009. Break it down: An alternative approach to measuring effectiveness in counterterrorism. Economics of Security Working Paper 23, Berlin: Economics of Security.

van Dongen, T. 2015. The science of Fighting Terrorism: The relation between terrorist actor type and counterterrorism effectiveness. PhD Thesis, Leiden University.

van Gulijk, C., S. Sillem, and M. Cayford. 30-Sep-2013. SURVEILLE Deliverable 3.4: Design of a research methodology for assessing the effectiveness of selected surveillance systems in delivering improved security. *Surveillance: Ethical Issues, Legal Limitation, and Efficiency Collaborative Project, Seventh Framework Programme*.

van Um, E., and D. Pisoiu. 2011. Effective counterterrorism: What have we learned so far? *Economics of Security Working Paper 55*, Berlin: Economics of Security.

Welsh, B. C., and D. P. Farrington. May 1, 2003. Effects of closed-circuit television on crime. *The ANNALS of the American Academy of Political and Social Science* 587(1):110–35.

Welsh, B. C., and D. P. Farrington. December 2009. Public area CCTV and crime prevention: An updated systematic review and meta−analysis. *Justice Quarterly* 26(4):716–45.

Willis, H. H., J. B. Predd, P. K. Davis, and W. P. Brown. 2010. Measuring the Effectiveness of Border Security Between Ports-of-Entry. Technical Report. RAND Corporation.

W&L Law Cybersurveillance Symposium Keynote: Gen. Michael Hayden, January 2015. https://www.youtube.com/watch?v= VUEuWiXMkBA.