The Good, the Cheap, and the Privacy-Friendly
Stakeholder Evaluation of the Effectiveness of Surveillance Technology

Cayford, Michelle

DOI
10.4233/uuid:63713a7e-b196-4bfc-a91e-e65bfad97fac

Publication date
2020

Document Version
Final published version

Citation (APA)
Cayford, M. (2020). *The Good, the Cheap, and the Privacy-Friendly: Stakeholder Evaluation of the Effectiveness of Surveillance Technology*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:63713a7e-b196-4bfc-a91e-e65bfad97fac

The Good, the Cheap, and the Privacy-Friendly

Stakeholder Evaluation of the Effectiveness of Surveillance Technology

Michelle Cayford

The Good,
the Cheap, and
the Privacy-Friendly:

Stakeholder Evaluation of the Effectiveness of Surveillance Technology

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof.dr.ir Tim van der Hagen
Chair of the Board for Doctorates
to be defended publicly on
4 June 2020 at
15 o'clock

By

Michelle CAYFORD

Master of Arts in International Security, Institut d'études politiques de Paris, France
born in Spokane, WA, USA

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof.dr.ir. Pieter van Gelder	Delft University of Technology, promotor
Prof.dr.ir. Genserik Reniers	Delft University of Technology, promotor
Dr.ir. Wolter Pieters	Delft University of Technology, promotor

Independent members:

Prof.dr. Michel van Eeten	Delft University of Technology
Prof.dr. John Mueller	Ohio State University
Prof.dr. Ibo van de Poel	Delft University of Technology
Ir. Ronald Prins	TIB (Toetsingcommissie Inzet Bevoegdheden van de inlichtingen-en veiligheidsdiensten)

SEVENTH FRAMEWORK PROGRAMME

This research has been funded in part by the European Union’s Seventh Framework Programme SURVEILLE project, under grant no. 284725.

ISBN: 978-94-028-2068-3

Cover design by Caleb Bringhurst

To my dear husband
By her who best knows his value

(dedication borrowed from Elizabeth Gaskell)

Table of Contents

List of tables and figures:

Summary

Watching without being seen and listening without being heard with the aim of gaining strategic advantage. This is surveillance and every nation that has the means to do so, engages in it. Government surveillance is intended to increase a nation's security; in so doing it inevitably invades others' privacy. Although surveillance is not a new phenomenon, and the debate over privacy and security is not novel, the course of this PhD has seen discussion renewed over these subjects as they have moved from the shadows to a topic of dinner table conversation (or perhaps, more accurately, to a topic of social media exchange). The debate typically focuses on privacy vs. security and whether or to what degree one should be exchanged for the other.

Fueling the current debate is not traditional, human surveillance, but surveillance technology. Surveillance technology includes a myriad of instruments, including CCTV, drones, satellites, listening devices, and perhaps particularly, wiretapping (of phones and internet) and technology that monitors and analyzes internet activity, such as deep packet inspection. The purpose of employing this surveillance technology is to gain strategic advantage over the adversary, whether that be a state actor or criminal or terrorist organization. Whether the technology contributes to gaining that advantage – i.e. achieves the security goal for which it was employed – is a question of effectiveness.

Although less discussed, effectiveness is a central element of the surveillance debate. To judge if a technology's surveillance is proportionate to the privacy invasion incurred, if or to what degree the technology contributes to increasing security must firstly be established. Evaluations of technology typically focus on performance – how well the technology operates, such as the quality of the image it produces. Effectiveness, however, examines whether gaining that footage contributes to the security goal of identifying members of a terrorist organization, learning about a state's nuclear program, etc.

Studies on the effectiveness of surveillance technology are scarce. Limited research has been done related to law enforcement use of surveillance technology and effectiveness, and concerning the effectiveness of counterterrorism programs. Apart from this, and CCTV research tackling the question of effectiveness in recent years, few studies on effectiveness have been carried out. This lack is particularly apparent in the realm of intelligence work, the domain largely concerned in the surveillance debate. This gap in research led to this dissertation's research question: *How is the effectiveness of surveillance technology evaluated in intelligence work?* The purpose of this PhD is not to measure the actual effectiveness of pieces of surveillance technology, but to examine stakeholders' treatment of this topic and the discourse surrounding it.

A serious consideration of effectiveness quickly leads to a consideration of the complexity of the issue and a realization that effectiveness is not determined in a vacuum. That is, determination to deploy or continue use of a technology rests on considerations of cost and proportionality, as well as whether it is strictly effective or not. This dissertation, therefore, in studying how effectiveness is evaluated, also analyzes how cost and proportionality are considered in this assessment. This is referred to as "overall effectiveness," which includes the elements of cost, proportionality, and "strict effectiveness."

Virtually everyone, whether conscious of it or not, has an interest in surveillance technology being effective. This includes the general public, law enforcement and intelligence officials who use the technology, privacy advocates, oversight bodies, and policy makers – the consumers of intelligence. This PhD examines how three of these stakeholder groups evaluate the effectiveness of surveillance technology in intelligence work: intelligence practitioners, oversight bodies, and the public. A separate chapter is devoted to each.

The individual studies of this dissertation are as follows:

The first study, in Chapter 2, lays a groundwork of examining several National Security Agency (NSA) surveillance technologies: wiretaps, PRISM, decryption, exploitation, and analysis tools and databases. This provides a general framework of some of the kinds of technologies employed by intelligence services. The focus on the NSA was a result of the center stage position forced upon this agency and its surveillance due to the Snowden leaks, which led to these types of technologies being at the center of the surveillance debate. These technologies were analyzed to understand how they function and what they do, and to what extent the label "mass surveillance," a categorization assigned by the media, was accurately applied. (Although the media used the term "mass surveillance" here we replace it with "bulk collection.")

This chapter cut through a lot of the noise surrounding NSA's surveillance technologies, to analyze their capabilities and classify them on a scale of targeted to bulk collection technology. The results inform the rest of the dissertation with an understanding of the kinds of technologies under discussion. On the one hand, they demonstrate that this discourse includes a wide range of technologies – from bulk collection to targeted, with various in-between possibilities. On the other hand, they help in understanding that when stakeholders debate questions of proportionality, they have technologies capable of bulk collection in mind.

Chapter 3 investigates what intelligence practitioners say regarding the effectiveness of surveillance technology. While it is impossible to fully know *how* practitioners evaluate, given the classified nature of their work, analysis of their statements regarding effectiveness sheds enough light on the subject to determine what measures of effectiveness they value. This study analyzed public statements made by the directors of the U.S.'s NSA and Central Intelligence Agency (CIA) and the U.K.'s Government Communications Headquarters (GCHQ) from 2006 and 2016, and conducted 8 interviews of officials involved in the output of intelligence.

Findings showed effectiveness evaluations to be lacking; they also unveiled, however, seven measures of effectiveness valued by intelligence officials. These measures were not identified from evaluations performed, but were dug out of practitioners' discourse on surveillance technology. The measures were: attacks thwarted, lives saved, criminal organizations destroyed, output (reports generated and data used in criminal cases), context (how surveillance programs complement other tools), support rendered to other agencies, and an informed policy-maker. While the subject of cost did not often appear in the data analyzed, cost considerations were found to be the driver behind formalized evaluations of surveillance programs. Effectiveness, strictly speaking, was rarely spoken of, but when it was addressed it was usually in the context of cost. Proportionality, on the other hand, was a much-discussed topic. Intelligence officials consider proportionality to be established by the law and oversight bodies; they themselves then act within these parameters. Significantly, practitioners make a sharp distinction between bulk data collected according to algorithms, and the limited selection of this data that is seen by human eyes.

Chapter 4 turns to oversight bodies and their evaluation of effectiveness. It studies the oversight bodies of the NSA, CIA, and GCHQ, analyzing the public documents and statements issued by these bodies from 2006 to 2016. This study found that like intelligence practitioners, oversight bodies rarely treat the question of the effectiveness of surveillance technology. Only one instance was found in which an oversight body directly addressed this issue. Rather, oversight bodies look to and call on intelligence agencies to perform such evaluation. Values and measures of effectiveness that oversight bodies hold mirror those of intelligence officials. These findings point to a (perhaps inevitable) dependency of oversight bodies on intelligence officials, who work with and know the surveillance technology, to indicate how to assess the technology and to carry out the evaluation. This chapter unearthed a trilemma of successfully evaluating the three elements of effectiveness, cost, and proportionality simultaneously. Oversight mechanisms are typically established to oversee one of these three elements, focusing on questions of budget or legality, for example. Even in cases where their mandate covers all three, they never evaluate all three simultaneously. These are three conflicting goals, making it an impossible trilemma for an oversight body to successfully address all three at once.

Chapter 5 studies the public as a stakeholder, using surveys of Dutch students and their parents over three consecutive years. The study was designed to examine the public's views on the effectiveness of surveillance technology, and the relation between its perceptions of effectiveness and views on privacy, and effectiveness and views on acceptable cost of surveillance technology. Findings indicated that the public does not engage in a trade-off with regard to effectiveness-privacy, nor with regard to effectiveness-cost, but rather expects all three – effectiveness, privacy, cost – to be delivered simultaneously. Supporting other studies, this work indicated that people tend to be divided into groups of being either trusting or concerned. Those that are trusting believe surveillance technology to be effective and non-privacy invasive, while those that are concerned find surveillance technology to be both ineffective and invasive of their privacy. This chapter found that these groups may be age-dependent, with students not yet solidified into being either uniformly trusting or concerned.

Chapter 6 revisited the findings in chapters 3 and 4 of intelligence practitioners and oversight bodies rarely addressing the question of effectiveness, to dig deeper and find the source of this lack. This study began with the hypothesis that bureaucracy may be a root cause, and used James Q. Wilson's work to examine this hypothesis. The results of this research found the hypothesis to be correct. Bureaucracy in democracies results in intelligence agencies having conflicting goals because the legislative bodies that direct them cannot decide which goal to give priority; in intelligence agencies being procedural organizations which means they are difficult to oversee and their outcomes difficult to evaluate; in oversight bodies establishing rules and procedures but not helping in assessing if the job has been done well; in workers and managers having an inclination toward what is easier to perform and measure (cost and legality vs. effectiveness); and in politics in a democracy resulting in government being bound to its constituents and their conflicting demands.

In answer to the main research question of how the effectiveness of surveillance technology is evaluated in intelligence work, this dissertation found that effectiveness evaluation is lacking. Only two instances of evaluation were identified and these were one-off assessments. While the research identified measures of effectiveness that intelligence practitioners and oversight bodies value, actual evaluation is largely absent. This lack can be attributed to bureaucratic constraints found in a democracy, which result in conflicting goals, employees concentrating on what is more easily measurable (e.g. cost rather than effectiveness), and intelligence agencies being difficult to oversee and with outcomes elusive to evaluation. How the public views effectiveness evaluation is diametrically opposed to the approach of government stakeholders. The public demands effective surveillance technology, at a reasonable cost, while protecting privacy, while the government recognizes an impossible trilemma of delivering all three of these factors simultaneously.

Democratic governments, thus, find themselves in an impossible situation of trying to simultaneously meet all the public's conflicting demands, while in reality being forced to perform trade-offs. As a possible way out of this impasse, this PhD proposes an independent third-party evaluation of overall effectiveness (strict effectiveness, cost, and proportionality). The reviewer would be highly qualified, possessing the necessary security clearances and having access to classified information, but the report itself would be fully public. This approach would bring neglected elements of effectiveness into the debate, as well as break the charged starting point of security vs. privacy. It would also provide more transparency, a necessary element for intelligence agencies to gain the public's trust.

This dissertation concludes with a reflection on additional actors in the effectiveness debate, and a discussion of frameworks and trust. Two frameworks for evaluating surveillance technology were developed in connection with this PhD; the dilemma of such an approach is discussed. That is, while assessment models may be useful in theory, if root causes of the lack of effectiveness evaluation are not examined and addressed, then creating evaluation models becomes an empty exercise. Such models will never be used if the root causes of the lack are not identified and dealt with. The issue of trust is also raised. Arguably, the core issue surrounding surveillance is not its effectiveness nor its invasiveness, but a matter of trust in the institutions conducting it.

Limitations to this research include the decision to focus on intelligence work, which is forcibly secret. Relying on open sources means the picture of intelligence agencies' and their oversight bodies' treatment of effectiveness is perhaps incomplete. For the study on the public, Dutch students and their parents were the subjects, with survey questions focusing on certain types and scenarios of surveillance technology. Different types of technology in different scenarios and/or another sector of the public may yield different results. Lastly, in order to facilitate research, slightly varying angles were taken on analyzing the various stakeholders: intelligence practitioners' statements on effectiveness, oversight bodies' evaluation of effectiveness, and public views on effectiveness.

This dissertation has scientific implications both within and outside the security field. For example, the security trilemma and its conflicting goals of effectiveness, cost, and proportionality could be applied to smart metering, commercial collection and use of online data, and health care monitoring systems – areas that are, in fact, surveillance of another kind. The finding of the lack of effectiveness evaluation could also be applied to other arenas (e.g. the private sector, environmental programs) to determine to what degree programs and technologies are assessed for actual effectiveness in other domains. The realities of bureaucratic constraints discussed in this dissertation are relevant for studies on government reforms, intelligence and oversight reforms, and questions of governance vis-à-vis surveillance technology. The identified measures of effectiveness valued by intelligence practitioners and oversight bodies are useful for future research in privacy and in policy. And the identified impasse created by public demands to deliver all three elements of effectiveness, cost, and proportionality simultaneously versus the reality of the security trilemma faced by the government could (and should) be incorporated into future surveillance studies.

Social implications include a proposal for an independent reviewer of overall effectiveness. More generally speaking, this dissertation provides a starting point for establishing objectified evaluations of effectiveness, and for a more honest dialogue concerning the complexities of effectiveness assessment, the security trilemma, and the fetters of bureaucracy, all of which directly influence surveillance technology and its evaluation.

Propositions

1) Intelligence agencies and oversight bodies face a security trilemma of being forced to give up one element of strict effectiveness, cost, or proportionality in order to successfully deliver the other two. *(This proposition pertains to this dissertation.)*

2) The surveillance dialogue needs to be more honest, discussing the complexities of evaluating effectiveness, the security trilemma, and bureaucratic constraints. *(This proposition pertains to this dissertation.)*

3) A possible partial remedy to the lack of effectiveness evaluation is an independent reviewer possessing the necessary security clearances and writing a public report. *(This proposition pertains to this dissertation.)*

4) Democracy in the age of social media may not prove able to withstand the subversive tactics of non-democracies intended to foment, deepen, and solidify divides among the population.

5) Much of the surveillance debate will continue to focus on privacy vs. security without much focus on effectiveness, as politicians and advocates seek out politically charged and divisive topics that can be easily used for political agendas.

6) The core issue of debates surrounding government is a matter of trust. No amount of new policies and laws will quiet critics distrustful of government bodies.

7) The secret sauce of private surveillance – what is collected, how and how long it is stored, etc. – is arguably more worrisome than government surveillance, as the government is restricted at least to some degree by law, while the private sector is left largely unfettered by current regulation.

8) A democratic state's attention to its values – e.g. liberty of the people, limitation of state power – weakens its surveillance capabilities causing it to lose to its non-democratic, often dictatorial enemies who wield unrestricted surveillance to strengthen state power.

9) When the Dutch hit on a good thing they stick with it: Gouda cheese, Speculaas cookies, blue Smelne boats with rope trim, exactly 10 thesis propositions, etc.

10) Raising children is far more challenging, time consuming, sleep-depriving, and brain-rackingly difficult than any PhD, and yet no diploma is awarded.

Chapter 1
Introduction

1.1 Problem definition

The surveillance debate

Surveillance is not a recent phenomenon. Human beings have engaged in spying on each other since time immemorial, with examples dating as far back as Biblical times (The Bible, Numbers chp. 12). Surveillance provides a window into what the enemy is doing without the enemy knowing. This is crucial for military battles – without an intelligence edge, commanders are at great disadvantage in war. Sun Tzu, a Chinese general and military strategist from the 6th century, famously said, "Know thy enemy." The way to know one's enemy is to engage in covert activities. Knowing what the other is doing, allows one to counter his moves, prevent him from gaining the upper hand, and thereby advance one's own interest. Surveillance, obviously, is not limited to classic war scenarios. It is necessary for gaining strategic advantage in peace as well. Used for gaining insight on countries, groups, or individuals considered to be threats, or for gaining economic or political edge in negotiations, peacetime surveillance is an integral part of any country's security.

Although perhaps not as ancient, the debate that surrounds surveillance is also not new. This discourse turns around methods and the extent to which a democratic government conducts surveillance. This debate was renewed and intensified by the Snowden leaks, beginning in June 2013. Whether or not a government should conduct surveillance on its own citizens, whether non-citizens are fair game, collection parameters, and privacy protections are all part of the discussion.

At the center lies the question of what many call privacy vs. security. That is, whether or to what extent privacy must be exchanged for greater security. There are those who argue that a certain amount of intrusion is paramount to protecting a nation's security. And there are those who argue that privacy must come first and not be sacrificed for so-called security, or further still, that governments exaggerate the terrorism threat and the role of surveillance to justify their use of surveillance. Literature on the subject abounds, discussing it from a plethora of angles and perspectives. At the heart of surveillance and, therefore, of this debate sits surveillance technology.

The technology

While the traditional method of simply observing with human eyes still exists, the use of technology is central to modern surveillance. Indeed, it is the surveillance via technology that seems to fuel the debate. These technologies include a vast spectrum, with video surveillance, such as CCTV, being perhaps the most prominent. CCTV comes in a multitude of forms, such as stand-alone or part of a modular system, activation-triggered or on continuously, detection of objects or of motion, fixed position or rotational, infrared, etc. Other types of technology, such as image detection systems, are often used with the CCTV for purposes like traffic monitoring. This means the system automatically detects certain activity (e.g. congestion) without human monitoring.

Satellites and drones with mounted cameras also perform surveillance through imagery. These are most often associated with surveillance of foreign countries, although in recent years, controversy has arisen over police department use of drones.

Audio surveillance in various forms of “bugs,” or listening devices (perhaps bringing to mind James Bond style gadgets) can be placed in cars, homes, conference rooms, public spaces, or any manner of object (e.g. a coffee cup) to listen in on conversations. Wiretaps are another method of surveillance through listening. This involves surveillance officers tapping into a suspect’s phone calls to gather evidence of criminal activity.

In the modern era of computer and internet activity, surveillance technology can include key logger software, which logs everything that is typed on a specific computer, as well as technology that monitors internet activity. Wiretaps exist in this arena as well, capturing data off fiber-optic cables. Deep packet inspection (DPI) monitors and analyzes internet traffic in real time. It is configured to identify certain kinds of traffic, such as specific IP addresses or VPN connections. All these surveillance technologies, from CCTV to DPI, are intended to contribute to gaining a strategic advantage through some form of “seeing” and listening.

The evaluation

Whether or not these technologies actually serve to gaining the sought-for advantage, or put another away, whether they accomplish their intended security goal, is a matter of effectiveness. The security goal could be learning the intentions of a country’s leader, the level of its military force, the hierarchy of a criminal organization, the members of a terrorist cell, or the location of a wanted person’s hideaway. Whether or not the surveillance technology contributes to achieving this goal is a starting point for determining if the technology should continue to be used for this purpose.

The question of effectiveness is strongly tied to the surveillance debate. The focus of this debate typically centers on privacy, but determining to what degree the technology contributes to the security goal is an equally, if not more, fundamental starting point. To judge if the privacy-security exchange is proportionate, one must first establish that the surveillance technology is contributing to the security side of the equation. This requires an evaluation.

Evaluation of surveillance technologies typically brings to mind performance-oriented assessments, rather than effectiveness evaluations. Performance focuses on how well a technology operates, while effectiveness refers to whether the technology contributes to the security goal of identifying members of the criminal organization, or whatever the case may be. The performance of a technology – e.g. image or sound quality – is obviously necessary for good surveillance. Studies evaluating surveillance technology tend to be technical in nature or be compartmentalized to one kind of technology. For example, Venetianer and Deng (2010) discuss challenges and solutions to evaluating intelligent video surveillance technology. The authors also describe the development of performance evaluation in the computer vision community, referencing numerous papers that evaluate vision algorithms and that discuss testing requirements. Other papers discuss tracking-based event detection, specifically in relation to traffic and transportation (Buch et al. 2009, Fuentes and Velastin 2004). These examples highlight evaluation of surveillance technology performance – how often the system accurately detects congestion, classifies vehicles, etc.

In the realm of CCTV, studies have moved beyond only assessing performance to determining whether or not the technology is effective in carrying out its intended security goal. Initial studies gave mixed and contradictory results, but more recently CCTV has been found to be effective in certain environments (Gill and Spriggs 2005, Ratcliffe 2006, Welsh and Farrington 2009, Caplan et al. 2011).

Among studies on surveillance technology effectiveness, the domain of CCTV is the most developed. Other research on the topic focuses on law enforcement use of surveillance technology, rather than on intelligence agencies. This is likely due to law enforcement work not being classified and therefore, more readily accessible. As a whole, however, apart from CCTV, a great deal of effectiveness studies do not exist.

The stakeholders

Whether these technologies serve their purpose and are effective is of interest to everybody. The "everybody" that has a stake in surveillance technology being effective includes practitioners who use the technology, both in law enforcement and intelligence; government officials on the receiving end of intelligence; oversight bodies; privacy advocates; and the public.

In law enforcement, officers, detectives, and analysts make use of surveillance technology to investigate crime, to gather evidence against and ultimately prosecute suspects. Without effective surveillance technology law enforcement officials cannot successfully perform their job. Likewise, intelligence practitioners depend on effective surveillance technology to gather information on state and non-state actors. Intelligence analysts and operations officers rely on this technology to deliver intelligence to heads-of-state and functionaries. The government officials consuming the intelligence make decisions on policy and action to be taken against states, terrorists, criminal networks and individuals, etc. based on the intelligence they receive.

Oversight bodies evaluate the performance of intelligence services and ensure these agencies stay within the bounds of the law in carrying out their duties, and that they do not waste taxpayer money in doing so. Surveillance technology is a factor in all these areas. Oversight bodies carry the power to penalize the intelligence services and to restrict and govern through law and regulation how the agencies carry out their job.

Privacy advocates champion individuals' right to privacy. Consequently, they often find themselves at odds with government officials over the use of surveillance technology. Privacy advocates are equally in favor of effective surveillance, but not at the expense of privacy. In their view, the protection of innocent citizens' right to privacy comes first and should be protected at all costs. Perhaps more than any other party, they challenge whether this technology is, in fact, effective.

The public's interest in effective surveillance technology comes at every level as its security and national interests, tax money, and privacy are all affected by it. Public opinion in democracies carries great weight with the public holding the power to elect and dethrone representative members of government, ultimately influencing the enactment of law and policy.

Effectiveness evaluation and its complexities

Effectiveness and its evaluation is a complex issue, as a serious discussion of the topic quickly reveals. This can be demonstrated with the well-known example of former NSA director General Alexander's statement in 2013 that certain NSA surveillance programs (then under attack due to the Snowden leaks) had disrupted 54 terrorist activities. While the accuracy of this number was much debated in the media, consideration of this case should lead to the deeper point of how does one measure effectiveness. The number of attacks thwarted, plots disrupted, or would-be terrorists arrested might seem an obvious measurement of success. But this only applies to terrorism cases and to technology that is yielding information directly related to an attack. Surveillance is used for much more than counterterrorism, and in the world of intelligence much of surveillance is used for informing policy makers and heads of state. Information comes from various sources to form an (incomplete) whole. Some surveillance technology may yield background information about a situation or group, but not the "golden" key information sought. In these cases, should effectiveness be assessed according to the information gained via the technology? This raises the obvious questions of quality and quantity. A technology should not be deemed effective just because long-winded reports are produced if those reports are of little value. But it is also not so clear that a technology should continue to be used if it has yielded only one piece of useful information over an extended period of time, even if that information was key. Presumably, most surveillance technologies will be effective some percentage of the time, but where does the threshold lie that establishes a technology as effective? Does it need to be effective 50% of the time? More? Less?

Considering these questions reveals the complexity of assessing effectiveness. It is not a simple question of counting the number of attacks prevented. Who is doing the assessing also plays a role. What might be considered effective by one party is not to another. On the U.S. Privacy and Civil Liberties Board, a majority of three members found a particular surveillance program to be ineffective, while one member judged it to be effective based on the information it collected being potentially useful in a future time-sensitive scenario. In other words, the surveillance program had collected information that currently had shown no usefulness, but in a future scenario of an impending terrorist attack this information could be searched, potentially yielding key information to stop the attack. In the view of this board member the potential usefulness of this information rendered its collection and thereby the surveillance program employed, effective. The fifth board member found the program to be effective based on its use to triage threats and verify intelligence collected elsewhere.

Both of the above dissenting board members judged the surveillance program in question to be effective, not based solely on the usefulness of the program, but also in consideration of privacy intrusions. They concluded that the program was effective based on the value they attributed to the program coupled with what they believe to be a small actual privacy intrusion (due to no location or personal identifying information being collected, as well as to the strict safeguards in place regarding use of the data). As evidenced here, effectiveness is not determined in a vacuum. Even once measures of effectiveness have been established, the effectiveness evaluation does not stop there.

Considerations of proportionality play a role in ultimate determinations of effectiveness. Likewise, cost is a very real practical factor that influences decisions to deploy or continue use of specific technologies. These additional elements factor into stakeholder evaluations of effectiveness and ultimate effectiveness conclusions.

1.2 Research question

While the privacy topic has been well-studied and some understanding of law enforcement use of surveillance technology has been gained, effectiveness and surveillance technology use in intelligence is still poorly understood. This dissertation focuses on three elements related to the surveillance debate and to surveillance technology that have been minimally studied: effectiveness, surveillance technology, and intelligence. It brings together the themes of the privacy-security debate, the evaluation of surveillance technology, and effectiveness, examining what this dissertation calls "overall effectiveness" (effectiveness, cost, and proportionality) in the context of surveillance technology use in intelligence agencies. The resulting main research question is as follows: *How is the effectiveness of surveillance technology evaluated in intelligence work?*

This question moves from solely focusing on privacy or concentrating uniquely on technology to focusing on effectiveness, bringing these elements together as a greater whole. And it does so in the largely hidden realm of intelligence work. This dissertation examines how different stakeholders evaluate effectiveness and how they assess cost and proportionality in their evaluation.

Stakeholders

This dissertation examines three groups of stakeholders – intelligence practitioners, oversight bodies, and the public – and their treatment of effectiveness. The decision was made to focus on surveillance technology in intelligence agencies as opposed to law enforcement. Intelligence agencies are necessarily more secret and consequently less is known about them than police agencies. This renders them both more difficult and more interesting to study. The lack of study surrounding intelligence agencies, their use of surveillance technology, and their treatment of effectiveness, as well as surveillance by intelligence services being a front and center topic at the start of this PhD led to intelligence practitioners being chosen as a stakeholder.

Oversight bodies play an important role in "seeing" inside intelligence agencies and ensuring they act according to the law, and yet relatively little is known regarding how oversight bodies approach the topic of effectiveness. It seemed logical and necessary to follow the chapter on intelligence practitioners with a study of oversight bodies as a stakeholder.

The public was selected as the final stakeholder due to its ultimate influence on policy-making and laws surrounding surveillance. Few studies on the public have zeroed in on the effectiveness question and its relationship vis-à-vis proportionality and cost. Obviously, any study of the public is limited to a certain number or group. The sector of the public chosen here was Dutch university students studying security and their parents. This allowed for study of 1) the Dutch public, previously not done in relation to surveillance and effectiveness and 2) security studies students, also previously unexamined and potentially holding differing views to the rest of the population.

Surveillance technologies studied

The surveillance technologies discussed in this dissertation varies by chapter. In chapter 2 specific NSA surveillance technologies are named and analyzed. These technologies were selected prior to the study based on the media debate surrounding them and the availability of information on them. In the subsequent two chapters on intelligence practitioners and oversight bodies, “surveillance technology,” in principle, includes any and all types. These studies were performed without selecting specific kinds of technology, but rather analyzing data in which these stakeholders were discussing any kind of surveillance technology. Out of the data, technology dealing with communications (telephone data, internet, email, etc.) became the focus, as these were the kinds of technology under discussion. Other technologies, such as drones and satellites were also mentioned, but the vast majority of the material was related to surveillance technology treating communications data. Therefore, based on the data, these kinds of technology became the focus for these chapters. The study on the public was based on a survey, which asked questions about certain kinds of surveillance technology, making these types the focus of that chapter.

1.3 Methodology

The dissertation begins with an examination of several specific NSA surveillance technologies. These technologies were selected due to the Snowden leaks, which brought these kind of systems to the forefront of the surveillance debate. Thus, chapter 2 examines NSA technology that frequently appeared in news headlines. The choice of the NSA and its technologies versus another intelligence service was for the same reason – these surveillance programs were often under discussion, but without much analysis of the technology itself. Additionally, there was more information publicly available on these systems. This chapter answers the sub-question: *#1 What kinds of surveillance* *technology does the NSA employ and how can they be categorized?* The kinds of technology that became the focus for the chapters on intelligence practitioners and oversight bodies were also the types examined in this first chapter. Knowing the technology itself, how it functions and how it can be employed, is a fundamental starting point to any evaluation.

In chapters 3-5 the main research question – *How is the effectiveness of surveillance technology evaluated in intelligence work?* – was applied to the three groups of stakeholders of intelligence practitioners, oversight bodies, and the public. Due to the nature of these groups, slightly different angles were needed to analyze this question. Regarding intelligence practitioners, it is impossible to know how they evaluate effectiveness. Their work is classified – speaking about it is illegal and any pertinent documents are inaccessible. What *is* possible to study, however, is how intelligence officials speak about effectiveness in public settings and documents, thus allowing some insight into how they treat this subject. This method was employed, and the related sub-question was: *#2 What do intelligence practitioners say about the effectiveness of surveillance technology?* Oversight bodies, on the other hand, produce many public documents, making it more feasible to study how they evaluate effectiveness. Therefore, to study oversight bodies the documents they produce were examined to respond to the following sub-question: *#3 How do oversight bodies evaluate the effectiveness of surveillance technology?*

The public, of course, does not perform actual evaluations of effectiveness. It does not have access to the results, or intelligence gained, via surveillance technology, and therefore cannot measure its success. It does hold perceptions, however, of whether surveillance technology is effective or not. These perceptions were studied via surveys in response to the question, *#4 How does the public perceive the effectiveness of surveillance technology?* For all three stakeholders the following sub-question was also asked: *#5 When considering effectiveness, how are factors, such as cost and proportionality taken into account?*

A significant finding from the chapters on intelligence practitioners and oversight bodies is that there is a significant lack of effectiveness evaluation of surveillance technology. This led to asking, *#6 What is the underlying cause for the lack of evaluation of effectiveness?* Chapter 6 hypothesizes that bureaucracy might be a root cause of this lack and examines this question using James Q. Wilson's seminal work on bureaucracy.

In the conclusion, the stakeholder perspectives are brought together to answer the main research question. The impasse resulting from opposing perspectives and the security trilemma are highlighted, and reflections given on the utility of prescriptive evaluation models and the role trust plays in effectiveness judgments.

1.4 Dissertation outline

Table 1.1 below outlines the subsequent chapters of this dissertation and their associated publications. Chapter 2 analyzes and classifies specific NSA surveillance technologies. Chapters 3-5 deal with intelligence officials, oversight bodies, and the public, respectively. The final chapter concerns bureaucracy and its role vis-à-vis the evaluation of effectiveness.

Chp.	Title	Publication
2	NSA Surveillance Technology Explained	Cayford, Michelle, Coen van Gulijk, and P.H.A.J.M. van Gelder. "All swept up: An initial classification of NSA surveillance technology," In *Safety and Reliability: Methodology and Applications,* edited by T. Nowakowski, M. Młyńczak, A. Jodejko-Pietruczuk and S. Werbińska-Wojchiechowska, 643-650. CRC Press/Balkema, European Safety and Reliability Conference, Wroclaw, Poland, 2014.
3	What Intelligence Officials are Saying…	Cayford, Michelle and Wolter Pieters. "The effectiveness of surveillance technology: What intelligence officials are saying," *The Information Society*, 34:2, 88-103. 2018. DOI: 10.1080/01972243.2017.1414721
4	Oversight Bodies – Evaluation through Compartmentalization	Cayford, Michelle, Wolter Pieters, and Constant Hijzen. "Plots, murders, and money: oversight bodies evaluating the effectiveness of surveillance technology," *Intelligence and National Security*, 33:7, 999-1021. 2018. DOI: 10.1080/02684527.2018.1487159
5	The Public Wants It All	Cayford, Michelle, Wolter Pieters, and P.H.A.J.M. van Gelder. "Wanting it all – public perceptions of the effectiveness, cost, and privacy of surveillance technology." *Journal of Information, Communication, and Ethics in Society,* 18:1, 10-27. 2019. DOI: https://doi.org/10.1108/JICES-11-2018-0087
6	Fettered by Bureaucracy	Accepted for publication with *Intelligence and National Security*.
7	Conclusion	

Table 1.1. Dissertation overview

Chapter 2
NSA Surveillance Technology Explained[1]

2.1 Introduction

Many types of surveillance (phone tapping, collecting metadata, intercepting mail) used by the NSA are not new to the world of espionage. What is new in surveillance, thanks to huge advances in technology, is bulk collection. In the digital age, in which everyone's communication and information is stored and transmitted online it is now possible to gather massive amounts of information rather easily. What formerly would have required a laborious process of monitoring physical mail, tracking targets' movements, and entering homes or offices to gain information about a target's activities, contacts, financial information, etc., can now be done largely by monitoring their online activity. And in a post-9/11 era where terrorism is a primary concern of governments, all this information can be swept up not just about one person, but about thousands, potentially even everyone, in the event that this information may prove useful later, and in the interest of identifying and apprehending terrorists before they are able to act.

The objective of this paper is to gather information about NSA technical systems in order to form a basis of understanding of the technology from which to then discuss the effectiveness of surveillance technology. The analysis was performed from the perspective of risk control. That is to say, the assessment focuses on an integral analysis of the system rather than on computer science aspects.

[1] This chapter has been published as: Cayford, Michelle, Coen van Gulijk, and P.H.A.J.M. van Gelder. 2015. "All Swept up: An initial classification of NSA surveillance technology." *Safety and Reliability: Methodology and Applications.* Edited by T. Nowakowski et al., 643-650. CRC Press/Balkema, European Safety and Reliability Conference, Wroclaw, Poland. 2014.

The published article used the term "mass surveillance." Subsequent studies revealed, however, that a more accurate term would be "bulk collection." Strictly speaking, "mass surveillance" refers to the pervasive surveillance of an entire population or a substantial sector. This dissertation's findings reveal that the NSA does not engage in this kind of surveillance. Bulk collection, which as the name indicates, refers to collecting data in bulk, is a more accurate term. Therefore, here we replace the term "mass surveillance" with "bulk collection."

Other small edits have been made related to changes with follow-up studies.

Formatting changes have been made in this and all chapters for consistency.

This paper will provide an initial classification of NSA surveillance technologies as they have been revealed in the recent Snowden leaks. The classification aims to group the technology behind the confusing array of NSA operations and programs, and to rate it on a scale of bulk collection to targeted surveillance. The following surveillance technology categories are examined: wiretaps, PRISM, decryption, exploitation, and analysis tools and databases. We then classify these technologies as they relate to bulk collection or targeted surveillance.

Two factors have influenced our research. One is that what kind of surveillance technology the NSA uses and how it works is all classified information. Therefore, much of our examination and explanation is educated guessing based on recent leaks, previous revelations, and an understanding of what methods and devices are available. Secondly, the stories on the Snowden leaks have been published in a multitude of media sources in various countries. This renders it more difficult to gather pertinent information into a meaningful ensemble. Further, these leaks continue and information on these systems may consequently change. Therefore, these are initial steps to understand NSA surveillance technology and will be updated and improved as necessary.

2.2 NSA Surveillance Technologies

2.2.1 Wiretaps

Wiretaps on fiber-optic cables are likely the means by which the NSA accesses and captures vast amounts of internet data. This would provide the basis for it to analyze and map the information, identify specific targets, and deploy malware attacks. The conclusion that the NSA is tapping fiber-optic cables is based on three things: 1) According to an NSA slide, “Upstream” is “the collection of communications on fiber cables and infrastructure as data flows past.” 2) The GCHQ operation TEMPORA is known to be tapping into fiber-optic cables and to be working with the following telecom companies: BT, Verizon Business, Vodafone Cable, Global Crossing, Level 3, Viatel, and Interoute (Ball et al. 2013). 3) A 2006 court case against AT&T, which disclosed that the NSA was wire-tapping fiber-optic cables at AT&T’s internet exchange point in San Francisco.

Fiber-optic cables

Fiber-optic cables carry the world’s communications across the globe. More than 550,000 miles of these undersea cables connect our world (Fulghum 2010), carrying 99 percent of the world’s inter-continental data (Timberg and Nakashima 2013).

Standard fiber-optic cables over land consist of 144 individual glass fibers; those undersea consist of only 8 individual glass fibers. Each fiber carries 40-160 light wave signals and each of these signals handles 10-40 gigabits of traffic (Fulghum 2010). The data is turned into "ultra-short flashes of light. These flashes represent the zeros and ones that all digital information is comprised of. A photodiode at the end of the cable turns the light flashes back into electrical signals" (Schmidt 2013). The signals, however, have to be re-amplified with a regenerator about every 80 kilometers otherwise they will drop. Each fiber optic has to be to re-amplified separately, so the fiber optics must be laid out separately rather than bundled together. If a cable was to be secretly tapped this is the weak point at which it could be more easily done (Schmidt 2013).

Fiber-optic splitters

An even easier way to tap fiber-optic cables is to have the consent of the telecommunications company to split the wires. This was the case with AT&T – the company allowed the NSA to tap its cables.

To intercept AT&T's signal at 611 Folsom St. in San Francisco, the NSA installed splitters. A splitter is a piece of equipment that physically splits the cable and sends the signal in two different directions. The simplest form of splitter is a "T" which has one signal coming in and two signals going out. The NSA splitters used in this case were 50/50 splitters – of the signal that came in, 50% of it went out one fiber and 50% went out the other. This does not refer to 50% of the data, but to 50% of the signal. That is, the signal is split in half, making it weaker, but 100% of the data is sent through each of the two signals. Essentially this means that a copy of the data is being made. The data is then sent on its way to its original destination through one signal, and a copy of this data is sent into another cable owned and operated by the NSA (see Figure 2.1).

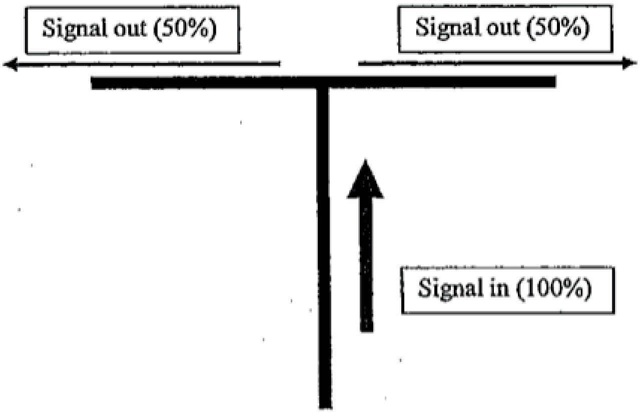


Figure 2.1. 50/50 Splitter (Source: Declaration of J. Scott Marcus, p.13)

Effectively, these splitters were installed at an Internet Exchange Point. At the San Francisco facility AT&T exchanged internet traffic with 16 companies with which it had peering arrangements. The splitters diverted traffic related to AT&T's Common Backbone, the network that provides internet access (as opposed to telephone traffic). It is believed that all, or substantially all, of AT&T's peering traffic was diverted through the splitters, but that its own traffic was not. The splitters were placed on the fiber-optic cables belonging to these 16 internet service providers (ISPs) before they reached the AT&T Common Backbone. Thus, their traffic was copied, but AT&T's own traffic was not (Marcus 2006).

Deep packet inspection

After passing through the splitters, the copied data was sent into a secure room at the San Francisco facility where it was processed by Narus equipment.

Narus is deep packet inspection (DPI) technology. DPI is the inspection and analysis of internet traffic in real time. It extracts basic protocol information, such as source and destination IP addresses, as well as the deeper layers of the traffic, which consist of the actual content of the traffic.

DPI was designed by engineers for ISPs to optimize their networks. "The primary technical capability underlying DPI is the ability to recognize. DPI has been developed to detect, for example, applications, protocols, media content, viruses or data in a specific format, such as credit card numbers" (Asghari et al. 2012). Once data is recognized the ISP can manipulate it, for example, by blocking or prioritizing certain traffic. It can also carry out notification actions, such as "generating reports, alarms or billing incidents" (Asghari et al. 2012). Using DPI, ISPs can prioritize applications like VoIP to improve service or identify illegal downloads. In other words, ISPs use DPI to look at the internet activity of their customers and act upon it.

The Narus system at the AT&T San Francisco facility consisted of two parts – the Narus Semantic Traffic Analyzer (STA) 6400 and the Narus Logic Server. Part one of the system – the STA 6400 – monitors data packets for metadata that matches "key pairs," such as "a specific IP address or a range of IP addresses, a keyword within a Web browser request, or a pattern identifying a certain type of traffic such as a VPN or Tor connection" (Gallagher 2013). The matching packets are put into the second part of the system, the Narus Logic Server. This second part consists of analytic processing systems that re-assemble the network sessions of the matching packets, mine them for "metadata, file attachments, and other application data" then index and deposit the data into a database (Gallagher 2013). Narus is capable of processing huge volumes of data and of storing the traffic that it captures.

The amount of internet traffic that Narus can handle monitoring is directly related to how many rules have been loaded into the machine watching the metadata flow past (part 1 of the system). The more rules, or pre-configured filters, turned on, "the more compute power burned and memory consumed per packet, and the fewer packets that can be handled simultaneously" (Gallagher 2013). In the Narus system, if all the pre-configured filters are turned on then the system can monitor 12 gigabits out of 20 (on a two-way 10 gigabit Ethernet connection). To make the system more efficient and able to monitor more of the 20 gigabits flowing past, some of these filters have to be turned off. "In other words, to handle really big volumes of data and not miss anything with a traffic analyzer, you have to widen the scope of what you collect" (Gallagher 2013).

Also present at the AT&T facility was what was referred to as the "SG3 backbone." This was a private backbone network, separate from AT&T's backbone. It was believed that this indicated that after the Narus equipment collected and processed data of interest, the data was sent via the SG3 backbone to central locations for further analysis (Marcus 2006).

Possible additional locations of NSA splitters

AT&T staff revealed that besides San Francisco, NSA surveillance sites also existed in Atlanta, Los Angeles, San Diego, San Jose, and Seattle. Using criteria based on the assumption that the NSA would want to intercept the most data with the fewest number of splitter sites, researchers at the University of Toronto have identified another 12 cities that likely host NSA splitters. Of 1,319 U.S.-only internet data traffic routes in the researchers' database, only seven did not pass through one of these 18 cities suspected to have NSA surveillance sites. That means that with splitters at these locations the NSA could intercept 99% of U.S.-only traceroutes (Clement 2013).

Concluding remarks on fiber-optic cable tapping

The splitters at the AT&T San Francisco location were never known to be removed. The aforementioned NSA "Upstream" slide mentions several code names in relation to the Upstream program – FAIRVIEW, STORMBREW, BLARNEY, OAKSTAR. These code names reportedly refer to specific telecommunications companies such as AT&T (Cohn 2013). It seems likely that the NSA surveillance technology being employed in Upstream are splitters placed on the fiber-optic cables along with DPI technology. The program, MUSCULAR, in which the NSA exploits the data links connecting Yahoo and Google data centers, could also be using splitters and DPI to tap the fiber-optic cables connecting the data centers.

Deploying splitters on fiber-optic cables is clearly bulk collection since a copy of all data is made. And in this instance DPI would also qualify as bulk collection surveillance since the parameters are set wide to capture as much data as possible.

2.2.2 PRISM

The Guardian article revealing the PRISM program reported that this program gave the NSA direct access to the servers of major internet providers such as Google, Apple, Skype, and Yahoo. The slide speaks of PRISM "collection directly from the servers" of nine U.S. internet services providers. The interpretation by *The Guardian* and *The Washington Post* was that this meant these companies were collaborating with the NSA to give it a direct connection to their servers, to "unilaterally seize" all manner of communications from them (Greenwald and MacAskill 2013). This proved, however, to be erroneous.

The "direct access" described here is access to a particular foreign account through a court order for that particular account, not a wholesale sucking up of all the information on the company's users. The court order gives the NSA access to the targeted account as well as to the accounts it is in contact with. This follows in the same principle of a court order on a phone number yielding the phone numbers the targeted phone has communicated with (with the obvious difference of a Facebook page containing content, while phone metadata does not). The NSA and the attorney general serve a court order on one of these companies for one or more foreign accounts, say, for example, several Facebook accounts in Pakistan. Those accounts are then monitored and their activity is sent back to the NSA. This is where "direct access" fits in – the NSA has access to the account in real time (Soar 2013). Marc Ambinder, writing in *The Week*, speaks of a "mirror" of the accounts that the company somehow creates and only the NSA has access to. When the selected account is updated the Facebook server and the mirrored server are both updated in real-time. PRISM is the tool that allows the analyst to monitor and analyze this data and all the data on foreign targets provided to the NSA by internet companies in the U.S. (Ambinder 2013).

PRISM, then, is a targeted technology used to access court-ordered foreign internet accounts.

2.2.3 Decryption, or circumventing encryption

One of the NSA's core missions is to decipher and break codes. Interestingly, however, the leaks have shown that the majority of the NSA's decryption efforts do not involve actually breaking codes, but exploiting human elements and implementation software to circumvent encryption. Further, while bulk collection programs aid in identifying machines to exploit, the actual decryption – or circumventing of encryption – is done on a targeted, rather than bulk collection scale.

BULLRUN is the NSA's program dealing with defeating encryption. The Classification Guide for Project BULLRUN states that the NSA has "some capabilities" against encryption in HTTPS, VoIP, Secure Sockets Layer (SSL, used for online banking), VPNs, SSH, and Webmail. The Guide also makes explicit that "'capabilities against a technology' does not necessarily equate to decryption."

The NSA has various methods for defeating encryption – obtaining encryption keys, backdoors, influencing encryption standards, and brute force. To obtain encryption keys the agency issues court orders requiring companies to turn over keys; hacks into companies to obtain the keys; or secures companies' willing collaboration. The agency reportedly maintains of database of the keys it collects (Simonite Sept. 6, 2013). The NSA has also persuaded or coerced companies into installing backdoors into their security systems. It can then get around the encryption by using these backdoors.

The agency purportedly influences encryption key standards. This would probably be done through the National Institute for Standards and Technology (NIST) "which sets U.S. cryptography standards and is influential worldwide" (Simonite Sept. 9, 2013). The MIT Technology Review's IT editor for hardware and software argues that it is unlikely that the NSA compromised many of these most widely used standards because they were developed by open groups outside the U.S. The one standard it did play a significant role in developing is part of a cryptography toolkit – Suite B – used by the U.S. government and its contractors. "Introducing backdoors into that would seem counterproductive to the NSA" (Simonite Sept. 9, 2013).

The NSA also uses brute computer force to break weak encryption. Therefore, although there is not yet evidence that the NSA has cracked SSL, experts have long warned that the keys typically used with SSL are not long enough. A government agency or a large company with significant resources could break the 1,024 bits long keys that most sites use for SSL. Longer keys are necessary to protect against this kind of attack, but few companies use them (Google and Facebook are just this year switching to longer keys. Google's key will be 2,048 bit.) (Simonite Sept. 9, 2013).

The NSA capabilities under BULLRUN are interesting because they show the limits of its abilities. The agency cannot break encryption. Its attacks against encryption use must be done on a case-by-case basis. However, it also illustrates its ability (and potential) to bring more data under its surveillance by bypassing encryption. Obtaining each individual key or backdoor access is targeted; what this yields could be targeted or bulk collection surveillance (e.g. it could read all the traffic of a company once it has the encryption key). Where on the bulk collection-targeted scale decryption falls is therefore somewhat variable. We have chosen to place it towards the bulk collection end of the scale but not fully in this category since encryption keys and backdoors can give it access to non-target data, but the agency does not have the capability to unilaterally bypass all encryption.

2.2.4 Exploitation

To exploit – hack into – computers and other devices the NSA utilizes its secret servers and software and hardware implants.

Secret servers

The NSA's attacks using FOXACID are a good example of both its exploitation techniques and how it circumvents encryption. FOXACID is a Computer Network Exploitation system that matches potential targets with prepared attacks. It is a modular system that allows exploits to be changed if discovered and only launches certain attacks against certain targets. FOXACID is used to perform all kinds of attacks. One example is attacks against Tor users (Schneier Oct. 4, 2013).

Tor – The Onion Router – is an online anonymity network. It is a problem for law enforcement because criminals use it for communication and it makes identification of the user, and hence, the criminal, excessively more difficult. Tor works by routing data packets through multiple nodes, or relays, rather than taking the most direct path. Each relay only knows the relay the data came from and the next relay it is going to. To the websites visited, the location of the Tor user appears random.

To execute an attack, the first step is to identify Tor users on the internet. This is easy for the NSA to do because the characteristics that make Tor anonymous also make all Tor users look the same on the internet. However, the user's identity and location remain unknown. "The NSA creates 'fingerprints' that detect http requests from the Tor network to particular servers" (Schneier Oct. 4, 2013). The fingerprints are then put into a database and data analysis tools are used to sort through all the internet traffic the NSA monitors to identify Tor connections.

Once a Tor user is identified the NSA uses its secret servers on the internet backbone, codenamed QUANTUM, to redirect those users to other secret servers, codenamed FOXACID. Because QUANTUM servers are at key locations on the internet backbone, they can react faster than other websites and thus impersonate the website the user is wanting to access. They can respond to the request before the actual website can and they look the same as the actual website. The target's browser is thereby fooled into contacting the FOXACID server. These kinds of attacks are a sort of race and are hard for anyone besides the NSA to execute because they depend on "a privileged position on the internet backbone" (Schneier Oct. 4, 2013). This is known as a man-on-the-side attack.

"By the time the NSA tricks a target into visiting one of those servers, it already knows exactly who that target is, who wants him eavesdropped on, and the expected value of the data it hopes to receive" (Schneier Oct. 15, 2013). Based on this, FOXACID automatically determines which exploits are best to serve against the particular target. It performs a risk-benefit analysis, considering factors such as the technical sophistication of the target, the value gained from a successful attack, the risk of discovery, the value and rarity of the exploit, etc. (Schneier Oct. 15, 2013). In the case of Tor, FOXACID attacks through the Tor browser bundle (a group of programs designed to make the installation and use of Tor software easier), exploiting vulnerabilities in the Firefox web browser. Once an attack has been successfully executed, the infected computer calls back to the FOXACID server, which then further infects the computer, compromising it long-term and providing the NSA with ongoing information (Schneier Oct. 4, 2013).

Note that although the NSA can execute these attacks against Tor it cannot do so on a large scale. Nor can it do so on demand. These are targeted attacks against individual users, not blanket attacks against all Tor users. According to 2012 NSA PowerPoint slides, the agency "will never be able to de-anonymize all Tor users all the time." It has had success de-anonymizing "a **<u>very small fraction</u>** of Tor users" with manual analysis, but has had no success in de-anonymizing a Tor user on demand (bold in original document).

This is apparently the same kind of attack that the GCHQ used against Belgacom. When targeted employees visited their LinkedIn profiles a secret served responded with a fake page that infected their computer with malware (*Spiegel Online* staff 2013).

Software and hardware implants

In December 2013 *Der Spiegel* published a leaked document of a sort of catalog of NSA spy technology. It included both software and hardware implants for exploiting mobile phones and their networks, computers, routers, and servers, among others. To give a couple of examples: SOMBERKNAVE, a software implant for Windows XP, uses the computer's unused wireless device to contact the NSA remote operations center and makes the targeted device controllable remotely. When the target is using the wireless card, SOMBERKNAVE is inactive. RAGEMASTER is a hardware implant hidden in the monitor cable that allows the NSA to see whatever appears on the monitor of the targeted computer.

Like the attacks performed by the secret servers, these implants are targeted surveillance, being targeted at specific devices. In some cases, however, they could move along the scale towards bulk collection depending on how they are used. For example, the implants for professional routers could potentially gather information on non-targets, depending on whether the organization itself or someone within the organization using the router is the NSA's target. It has also been reported by some sources that at least one of these systems – NIGHTSTAND – has been used from drones to target certain geographical areas (Appelbaum 2013). For this reason, we classify exploitation on the targeted end of the scale, but not completely in the targeted category.

2.2.5 Analysis tools and databases[2]

NSA has a myriad of data analysis tools and databases to process the data it collects. XKEYSCORE is both an analytical tool and database. We have chosen to focus on this system as an example of NSA databases, as it is often mentioned in reporting on NSA leaks, and is at the same time, perhaps, the most complex and difficult to understand.

After the attacks of 9/11, the NSA needed a quick way to increase its internet surveillance. It accomplished this by purchasing "off-the-shelf" systems such as Narus. The result was that a lot of data was collected. The problem for the NSA was where to store it and how to get it there. "Even when you store just the cream skimmed off the top of the 129.6 terabytes per day that can be collected from a 10-gigabit network tap, you're still faced with at least tens of terabytes of data per tap that need to be written to a database" (Gallagher 2013). It was physically impossible for the NSA to get all that

[2] It is debatable whether analysis tools and databases classify as surveillance technology since they do not collect any data themselves. We included this category in our analysis due to it being much discussed in the media at the time.

information back to its central database. So it created XKEYSCORE. XKEYSCORE solves the problem by storing the data packets in local caches rather than sending it back to a central database. Since the advent of XKEYSCORE the agency can now store 3 days' worth of raw packet data and 30 days' worth of metadata in the local caches.

XKEYSCORE is *not* the system that actually captures the data via the fiber-optic cable tap. The system that captures data packets from NSA wire taps is code-named TURMOIL. XKEYSCORE processes the data that TURMOIL brings in.

It processes it by running plug-ins, analysis engines that look for specific content in the captured data packets. XKEYSCORE has plug-ins for email addresses, phone numbers, webmail and chat activity, and extracted files, among others. "For selected traffic, XKEYSCORE can also generate a full replay of a network session between two Internet addresses" (Gallagher 2013). The plug-ins extract the metadata from each internet session and index it into tables. XKEYSCORE can track, cross-index, and search any kind of metadata that can be extracted from an internet session – log ins, email addresses, the use of encryption, language use, IP address geolocation, etc.

There are approximately 150 XKEYSCORE sites around the world. These sites include wire-taps at telecommunications companies' peering sites (such as the AT&T case), systems connected to friendly foreign intelligence agencies' collection sites, and mid-ocean fiber-optic cable taps (executed by F6, the joint CIA-NSA Special Collections Service). Only information that is related to specific cases is sent back to the NSA's central database. The data in the local caches is available to analysts through federated search while it is being stored (Gallagher 2013).

To perform a search request an NSA analyst creates a query. This query is sent to all the XKEYSCORE sites. Analysts can search by hard selectors (e.g. email addresses) or soft selectors (e.g. language). So if an analyst does not have a hard selector, such as an email address or phone number for a known target, he/she can do a search for a category of information, such as all encrypted Word documents or all VPN (virtual private network) startups in a given country (NSA XKEYSCORE slides 2008). Any kind of query can be created as long as the plug-in exists. XKEYSCORE combines and returns all the responses to the query.

Using XKEYSCORE is apparently how the NSA can identify Tor users – this is a category that can be queried. Another category that can be searched is exploitable machines – "Show me all exploitable machines in country X" (NSA XKEYSCORE slides 2008). When the NSA's Tailored Access Operations (TAO) identifies a computer as a target it loads its fingerprints, or unique identifiers, into XKEYSCORE's fingerprint ID engine. In the case of Microsoft system crashes XKEYSCORE identifies these error reports and sends an automatic notification, enabling TAO to exploit the machine (*Spiegel* staff 2013).

Like Narus, XKEYSCORE performs best when it "goes shallow," that is, fewer filters are applied to determine which data packets are captured. This means that a lot of information is collected, including, undoubtedly, information unrelated to NSA targets. XKEYSCORE, therefore, classifies as bulk collection technology.

2.3 Findings

Based on information currently available on the above NSA surveillance technologies, we have given them an initial and basic rating on a bulk collection to targeted surveillance scale (see Figure 2.2).

The NSA's use of splitters on the fiber-optic cables of companies like AT&T is clearly bulk collection. This surveillance technology indiscriminately copies all data, sweeping up vast amounts of data belonging to non-targets.

PRISM, by contrast, is a targeted surveillance method and is therefore placed at the other end of the scale.

The NSA's decryption program is a little less clear cut. The NSA does not have the capability to unilaterally bypass all encryption. However, once it does have an encryption key or a backdoor, for example, it can have access to far more data than just that belonging to specific targets. For this reason, we have classified it as one step down from the bulk collection surveillance of wiretaps.

The surveillance technology used for exploitation is targeted. However, as previously stated, in some cases it may be aimed at groups of people, which include non-target subjects. Therefore, we have classified exploitation as targeted surveillance but not completely in the category of targeted.

Whether or not analysis tools and databases are bulk collection or targeted surveillance is dependent upon the data that is put into them. Examining all databases is beyond the scope of this paper. We have taken XKEYSCORE as an example and found that because it sorts and stores volumes of data obtained via bulk collection methods, it is also bulk collection technology.

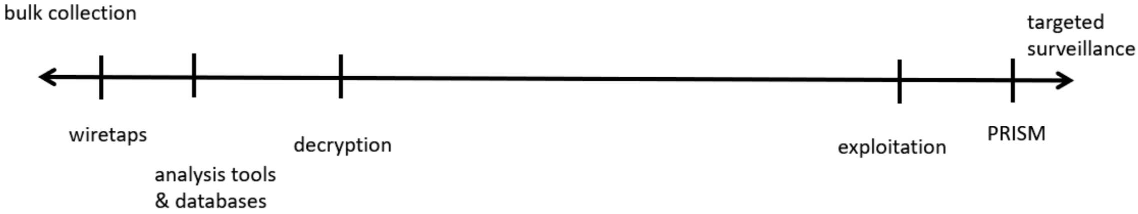


Figure 2.2. Classification of NSA surveillance technology

2.4 Conclusion

This paper aimed to create some structure for the many NSA surveillance programs recently revealed in the Snowden leaks. It was an initial step to classify the surveillance technologies and rate them according to bulk collection and targeted surveillance. We found that in most cases (PRISM being the exception) several programs could be classified under one technology category. Further, while some of these programs have been portrayed as bulk collection, they are, in fact, targeted. Lastly, further research is needed to determine if these technologies are effective.

Chapter 3
What Intelligence Practitioners are Saying…[1]

3.1 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.

[1] This chapter has been published as Cayford, Michelle and Wolter Pieters. "The Effectiveness of Surveillance Technology: What Intelligence Officials are Saying," *The Information Society*, 34:2, 88-103. 2018. DOI: 10.1080/01972243.2017.1414721

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. This 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.

3.2 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 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 and 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 and 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 and 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.

3.3 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.

3.4 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.

3.5 What Intelligence Officials are Saying About…

3.5.1 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" (Hayden in 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), "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" (p.14). 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).

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).

3.5.2 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.

3.5.3 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). 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[2] 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[3] 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). Although he judges the program to be effective, he does not give the metric by which he makes this determination. Changing

[2] Collection priorities are the policy issues and areas that policy makers task intelligence bodies to collect intelligence on (Lowenthal 2012, p.57).

[3] 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.

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 (Omand Oct. 2014, p.6). Here the use of data in criminal cases is seen as an indication of the effectiveness of the surveillance program that gathered the data.

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 the 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 3.1. 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.

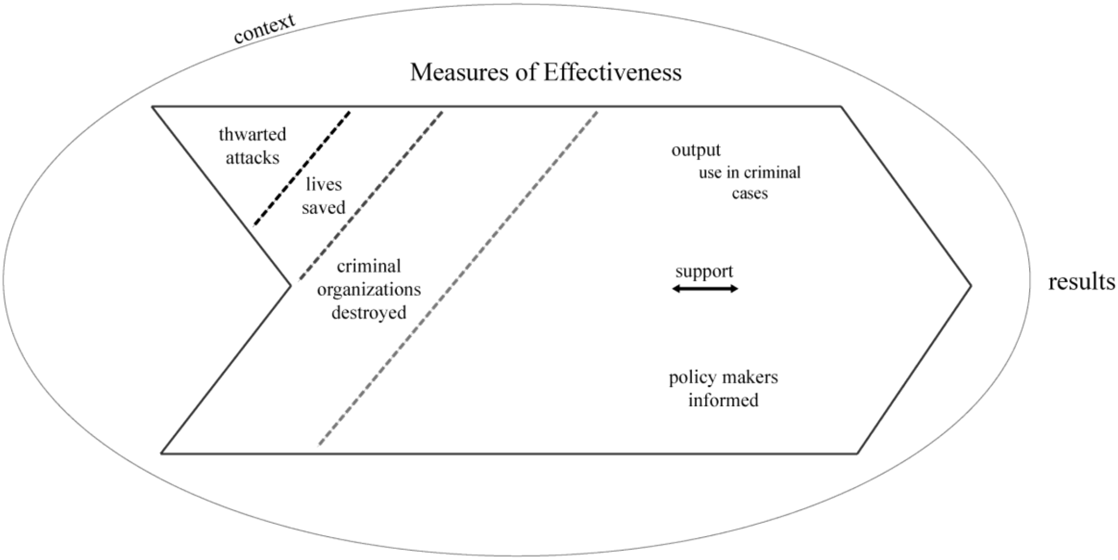


Figure 3.1. Evaluating effectiveness of surveillance technology as described by intelligence officials

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 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.

3.5.4 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). 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.

3.5.5 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).

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).

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"[4] 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

[4] "A selector must be a specific communications facility that is assessed to be used by the target, such as the target's email address or telephone number" (PCLOB July 2014, p.32).

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.

3.5.6 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[5] 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,

[5] 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.

"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 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.

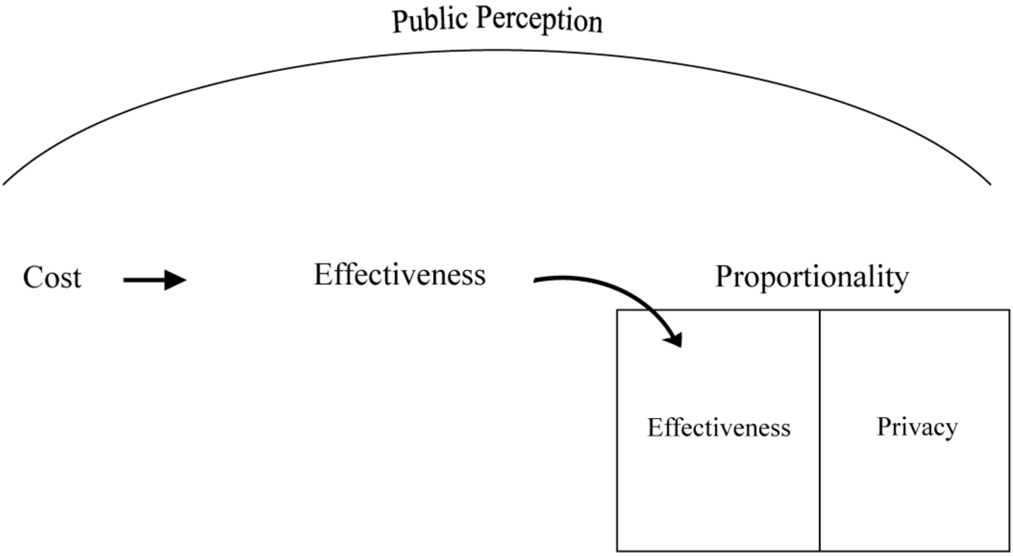
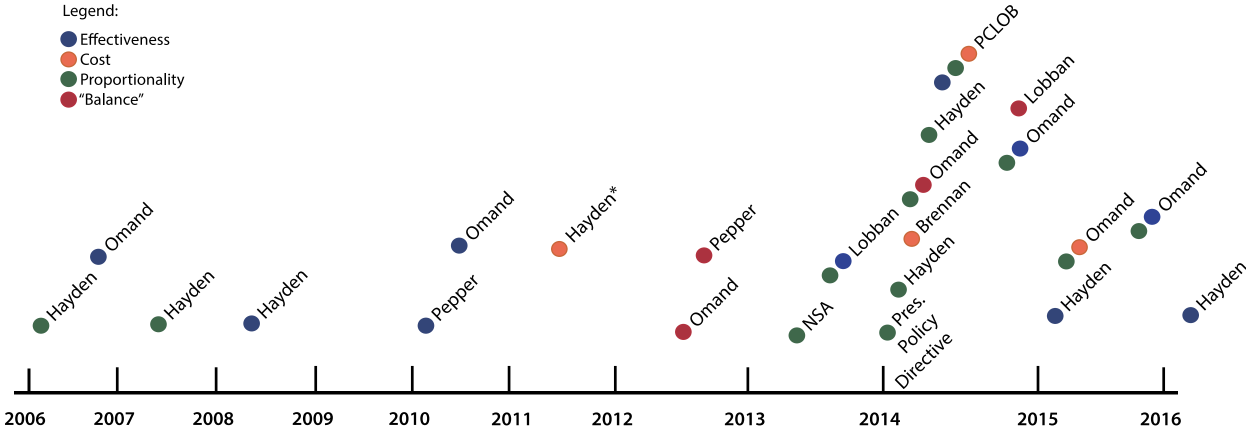


Figure 3.2. Components of overall effectiveness

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 3.2. 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.

3.6 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 3.3 plots over time intelligence officials' statements related to effectiveness, cost, proportionality, and "balance."[6]



*The original statement is from several years earlier than the 2011 article, but we were unable to determine the exact year.

Figure 3.3. Officials' statements related to effectiveness, cost, proportionality, and "balance"

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

[6] Statements made by interviewees are not included in this figure because the discussions were prompted by the author.

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.

Chapter 4
Oversight Bodies –
Evaluation through Compartmentalization[1]

4.1 Introduction

In September 2002 the U.S. Department of Defense received a Hotline complaint accusing the National Security Agency (NSA) of "fraud, waste, and abuse" related to the development of the TRAILBLAZER surveillance system, a data collection and processing program. The complaint alleged that the NSA had wasted money on TRAILBLAZER and had chosen TRAILBLAZER over the better (more effective) THINTHREAD program. As a result of this complaint, one of the oversight bodies for the NSA performed an audit on these two systems, concluding in favor of the complaint (Office of Inspector General 2004).

Recent years have seen an explosion of digital data. This has been followed by intelligence agencies trying to keep up with the flood of information in their endeavor to protect their countries against potential security threats. Rather than drown in the data, they strive to use it to more effectively identify and inform on security issues. As the flood of data rises, so too does the agencies' surveillance of that data.

Subsequently, the increase in surveillance provokes an increased concern of risk of abuse and privacy invasion – more surveillance powers to collect more data increases the likelihood that innocent persons' data may be swept up. To protect against this risk, governments often introduce increased oversight. A law introduced in the Netherlands in 2017 is a case in point. It gives more surveillance powers to intelligence services, allowing them to collect all the data traffic in a certain area in search of a terrorist; to hack the mobile phones of potential acquaintances of suspects; and to share collected, unanalyzed data with foreign intelligence services. To counterbalance this increase in powers, the law also introduces more oversight: the use of any of these new powers requires the prior permission of a special, new committee composed of two judges and a technical expert, in addition to the existing oversight of the Minister of the Interior and the Review Committee (CTIVD) (Kraan 2017).

[1] This chapter was published as Cayford, Michelle, Wolter Pieters, and Constant Hijzen. "Plots, murders, and money: oversight bodies evaluating the effectiveness of surveillance technology," *Intelligence and National Security*, 33;7, 999-1021. 2018. DOI: 10.1080/02684527.2018.1487159.

Both examples above demonstrate the key role oversight bodies play in intelligence agencies' use of surveillance technology. If this is the role (increasingly) given to oversight bodies, then how oversight bodies perform their evaluations becomes increasingly important. How, then, do these oversight mechanisms evaluate if a technology is effective in achieving its security goal? And how do they consider cost and proportionality in this evaluation? Do they consider all three in their oversight? Which, if any, of the three take priority? This comparative study investigates these questions, focusing on U.S. and U.K. oversight bodies, particularly those overseeing the NSA, CIA, and GCHQ. It is aimed at these two countries due to the Snowden leaks, which specifically focused on the surveillance of American and British intelligence agencies.

This study complements the authors' previous paper, which analyzed what intelligence officials of these agencies state regarding effectiveness (Cayford and Pieters 2018). It also compares the results of the previous study with what oversight bodies report on effectiveness, cost, and proportionality. Whether these two groups focus on the same things or different ones (e.g. counting money spent, plots thwarted, and murders averted) may impact audit results, like the one on TRAILBLAZER. Understanding how these groups consider the issue of effectiveness is an important first step for any subsequent dialogue on the use of surveillance technology and its governing laws.

This study is not an evaluation of whether or not surveillance technologies are effective; *nor is it a judgment* on how oversight bodies evaluate effectiveness. Rather, it is an examination, through analysis of public documents, of *how* oversight bodies deal with the question of effectiveness as part of their oversight function.

The next section of this paper presents related work in the fields of surveillance and oversight. The study's methods and approach are then presented, followed by an overview of the oversight mechanisms of the American and British intelligence communities. The data is then analyzed and the findings presented, followed by a discussion of these findings.

4.2 Related work

Two bodies of literature were reviewed for this paper: evaluations of the effectiveness of surveillance technology within the security domain, and existing studies on intelligence oversight.

Within the broad body of security and surveillance literature resides a set that deals with the strict effectiveness of surveillance technology. Strict effectiveness assesses and measures whether or not a given security program accomplishes its security goal. A significant body of works exists, which strives to evaluate the effectiveness of counterterrorism measures (Lum et al. 2006, Van Dongen 2009 & 2015, Van Um and Pisoiu 2011, Drakos and Giannakopoulos 2009, Jonas and Harper 2006). In addition, two government reports by the U.S. Privacy and Civil Liberties Oversight Board (PCLOB), address the question of strict effectiveness of security measures, revealing measures of effectiveness used by intelligence officials and drawing their own conclusions about the effectiveness of NSA surveillance programs (PCLOB Jan. 2014 & June 2014). In the U.K., David Anderson, an Independent Reviewer of Terrorism Legislation, reviewed the utility of the bulk powers used by the intelligence services (Anderson 2016). These powers include intercepting and acquiring telecommunications, equipment interference, and using personal datasets, all in bulk ("bulk" refers to large quantities of data, including those not associated with current targets). Anderson concluded that these powers are effective in achieving operational aims.

Mueller and Stewart (2011) evaluate the strict effectiveness of surveillance technology through cost-benefit analysis. Methods and frameworks related to strict effectiveness include Ekblom's work (2011), which establishes a framework for crime prevention and security in the community, and Sproles' (1999), which develops a method of establishing measures of effectiveness that can be applied to any field.

A small body of work deals with the strict effectiveness of specific kinds of surveillance technology, including assessing the effectiveness of U.S. Air Force drones, of border security (Lingel et al. 2012, Willis et al. 2010), of wiretapping programs (Tsvetovat and Carley 2006), and of advanced imaging technology full body scanners (Stewart and Mueller 2011). Lastly, there is an entire body of literature on CCTV cameras and their effect on crime. Certain recent studies have identified conditions in which CCTV operates most effectively, namely in small, defined spaces such as car parks, and against property crimes rather than violent crimes (Gill and Spriggs 2005, Caplan et al. 2011, Ratcliffe et al. 2009, Welsh and Farrington 2003 & 2009).

Our previous paper identified several measures that intelligence practitioners use to measure effectiveness: thwarted attacks, lives saved, criminal organizations destroyed, output, context, support, and informed policy-maker. The concept of effectiveness was found to be strongly linked to cost, with the goal of evaluating surveillance programs being efficient spending rather than effectiveness itself. Intelligence officials were found to rely on the law to determine proportionality and to consider collection by computer versus selection by human beings an important distinction. In the current paper, we compare these findings with those of oversight bodies.

Literature on intelligence oversight addresses the “basic problem” of “how to provide for democratic control of a governmental function and institutions which are essential to the survival and flourishing of the state but which must operate to a certain extent in justifiable secret” (Leigh 2007). This question has raised considerable academic attention within the intelligence studies since 1975, the American “year of intelligence,” when major scandals led to reforms in the oversight system (Johnson 2008). A “pattern of exposure, report, and strengthening of oversight” followed suit in many other countries throughout the 1990s (Leigh 2005).

Scholars have focused on the historical development of oversight, as well as performing comparative research (Barrett 2005, Johnson 2004, Kibbe 2010, Reid 2005, Snider 2008, Prados 2013, Johnson 2008, Schwartz 2007). Comparing oversight systems and practices in Argentina, Canada, Norway, Poland, South Africa, South Korea, the United Kingdom, and the United States, Born et al. (2005) concluded that factors such as independence from the executive, proper investigative powers, access to documents, the possibility to keep secrets, and sufficient support staff make oversight “strong.”

Much of this research focuses on legal, formal, and institutional factors that influence the institutionalization of oversight and control bodies (Born and Leigh 2005, Born and Caparini 2007, Willis 2007, Born et al. 2011, Willis and Vermeulen 2011, Born and Willis 2012, Born et al. 2015, Gill 2016). It describes different systems of oversight, ranging from parliamentary committees which exercise oversight on security and defense policies in general, to specialized committees, and non-parliamentary bodies. The research compares different aspects of these systems, such as their composition, selection of members, resources, mandates, the criteria they use, and temporal dimension (ex post or ex ante) (Willis and Vermeulen 2011).

Recently, several authors have started to research the cultural norms and social values that may influence intelligence and security practices, thus exploring the “soft side” of oversight and control (Davies 2004, O’Connell 2004, de Graff and Nyce 2016, Krieger 2009). Loch K. Johnson has shown how factors such as member motivation and cooperation by the executive influence the success of oversight and control in the field of intelligence and security (Johnson 2005). In the case of Dutch oversight it has been explored how informality and the lack of political weight characterize oversight practices (Hijzen 2014).

Much literature exists which examines the success of oversight bodies – how well do they perform their functions, are they formally equipped and actually using their powers to successfully oversee intelligence communities (Ford 2006, Ott 2003, Zegart and Quinn 2010, Dietrich 2016)? Aside from calls for modernization and “digitization” (Roy 2016), however, there is no literature on how oversight bodies assess whether intelligence agencies' use of surveillance technology is effective.

This current study works to fill this gap.

4.3 Framework and methods

Our conceptual framework relies on distinguishing different elements of effectiveness on the one hand, and the roles of technology, programs, and institutions on the other. This reflects the fact that the question of effectiveness is not determined in a vacuum. Other factors, such as cost and proportionality, are inevitably brought into consideration when determining whether or not to use a particular surveillance technology. We refer to this as overall effectiveness. Ultimately a decision on overall effectiveness includes considerations of strict effectiveness (whether or not the technology achieves the security goal), as well as of expense and proportionality.

This study defines *effective* as "an impact that is desirable and can be observed as contributing toward the sought-after security goal." Note that this differs from *performance*, which refers to the technology's ability to function correctly. For example, *performance* tells us whether a technology accurately captures targeted emails, while *effective* considers whether capturing those emails contributes toward dismantling the criminal organization.

Intelligence oversight is rarely properly defined – instead it is considered a catch-all term for all kinds of practices and institutions and used alongside terms such as "accountability," "review," and "control." However, most authors understand it along the lines of Hans Born's definition (Wegge 2017): "a means of ensuring public accountability for the decisions and actions of security and intelligence services" (Born et al. 2005, p.226).

In the documents analyzed for this study, the term "surveillance technology" as such, is not used. Rather, oversight bodies refer to "surveillance programs" or "collection programs." A program could entail only one surveillance technology, but more often it refers to several technologies that together perform a certain collection action, such as collecting internet data, filtering it, and selecting and storing the pertinent data. Consequently, this study often refers to surveillance programs rather than technology. This is with the understanding, however, that the programs focused on here are composed of technologies, and that considerations of effectiveness and the like can equally be applied to both. The types of surveillance programs being discussed in this paper are primarily those dealing with communications data (i.e. surveillance systems monitoring and collecting data on internet and phone activity). Figure 4.1 shows how the different elements of this framework fit together.

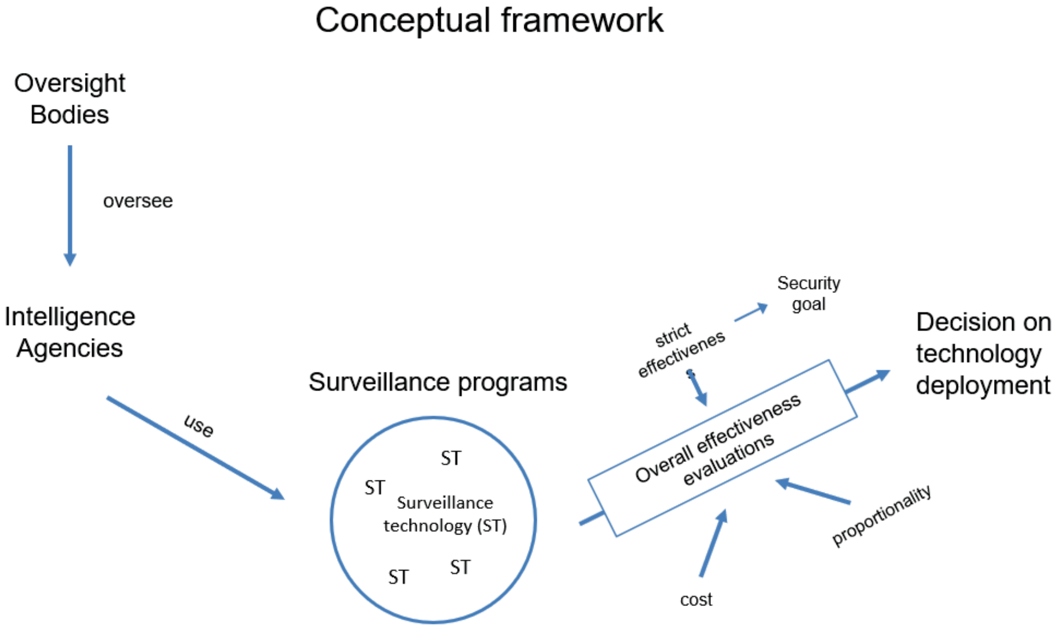


Figure 4.1. Conceptual framework of effectiveness, surveillance, and oversight

This study analyzes public documents and statements issued from 2006 to 2016 by the oversight bodies of the NSA and CIA in the U.S. and of the GCHQ in the U.K. These oversight mechanisms include the following: the British Intelligence and Security Committee (ISC), Investigatory Powers Tribunal (IPT), Interceptions Communications Commissioner, and Intelligence Services Commissioner; the American NSA and CIA General Counsels and Inspectors General, Director of National Intelligence (DNI), Attorney General, Department of Defense Office of Inspector General, NSA and CIA offices of privacy and civil liberties, Privacy and Civil Liberties Oversight Board (PCLOB), Foreign Intelligence Surveillance Court (FISC), and House of Representatives and Senate Select Committees on Intelligence. The timespan of 2006 to 2016 was chosen to provide a good amount of time (7 years) prior to the Snowden documents to potentially compare differences in evaluation pre- and post- Snowden. It is also the same timespan analyzed in the authors' paper on intelligence practitioners, allowing for possible comparison between the two studies.

The documents reviewed include all the documents available on the websites of the House and Senate Committees on Intelligence, excluding documents on their rules of procedure (47 documents reviewed). As concerns the FISC, all available – i.e. declassified – orders and opinions were analyzed (30). For the remaining U.S. oversight bodies, any available statements oral or written by these authorities were examined (27). One source – the audit on TRAILBLAZER – dating from 2004 was analyzed due to its extreme relevance. For the U.K., 13 ISC sources were reviewed, 10 reports by the Intelligence Services Commissioner, and 17 documents produced by the Interception of Communications Commissioner. All relevant judgments of the IPT were analyzed (7) – i.e. those in which the GCHQ was the defendant – as well as the two existing IPT reports.

Analysis was performed by identifying all mentions of effectiveness, cost, and proportionality. These statements were then evaluated, identifying measures of effectiveness and reoccurring themes in all three categories.

4.4 Overview of intelligence oversight bodies

Before presenting our findings, we briefly introduce the American and British oversight mechanisms and the bodies charged with overseeing the NSA, CIA, and GCHQ. The U.S. intelligence oversight mechanism is expansive and fairly complicated. Here we focus on the oversight bodies studied in this paper.

As the U.S. government has three branches – executive, legislative, and judicial – so does oversight. U.S. oversight can be considered as layers of an inverted pyramid. That is, the first layer of oversight is within the respective agency itself, and from there each successive layer fans out into increasingly broader layers (see Figure 4.2). There are several layers of oversight within the executive branch, including boards with particular oversight mandates, such as the PCLOB. The next layer of oversight is judicial – the Foreign Intelligence Surveillance Court (FISC) – but it is also partial, as its jurisdiction is limited to certain forms of investigative actions, such as electronic surveillance for foreign intelligence purposes. The final oversight layer is legislative and consists of the Permanent Select Committees on Intelligence in the House of Representatives and Senate.

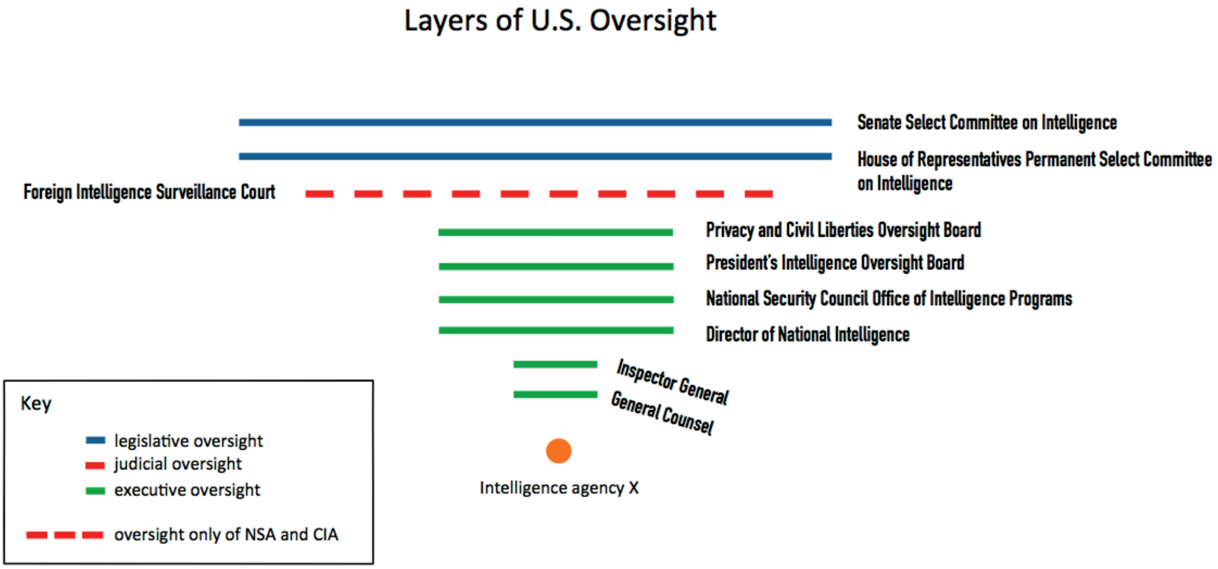

Layers of U.S. Oversight
Senate Select Committee on Intelligence
House of Representatives Permanent Select Committee on Intelligence
Foreign Intelligence Surveillance Court
Privacy and Civil Liberties Oversight Board
President's Intelligence Oversight Board
National Security Council Office of Intelligence Programs
Director of National Intelligence
Inspector General
General Counsel
Intelligence agency X
Key
legislative oversight
judicial oversight
executive oversight
oversight only of NSA and CIA

Figure 4.2. U.S. oversight layers

The British intelligence oversight mechanism consists of four pillars, which operate somewhat in successive order, from ministers to judicial commissioners, Parliament, and the Investigatory Powers Tribunal (IPT) (see Figure 4.3).

British oversight begins with ministerial oversight. For an intelligence agency to perform any given surveillance activity, a warrant or authorization is required. This authorization is given by certain designated ministers. The requested power is granted only if it is 1) legal, 2) necessary, and 3) proportionate. The second pillar is composed of Commissioners who review the agencies retrospectively, auditing their compliance with the law. The Intelligence Services Commissioner reviews all intrusive actions except interception; interception oversight falls to the Interception of Communications Commissioner. Under the new Investigatory Powers Act the Investigatory Powers Commissioner (IPC) will take over the responsibility of both the Intelligence Services and Interception of Communications Commissioners. Warrants will require the approval of both the IPC and a judicial commissioner. The Intelligence and Security Committee (ISC) of Parliament forms the third pillar. The ISC oversees the administration, policy, spending, and activities of the intelligence bodies. Lastly, the IPT exists for individuals to make complaint against the intelligence agencies if they believe they have been the victim of unlawful action or human rights infringement. The Tribunal then investigates and rules whether or not the complaint is justified, issuing orders for the agency in question, if necessary (IPT website).

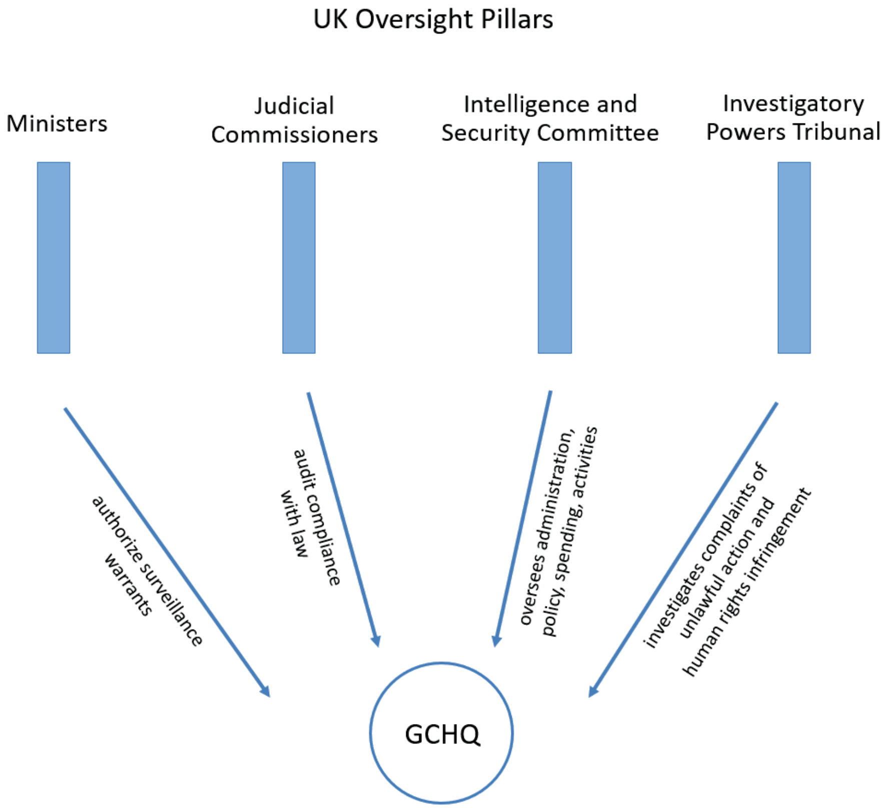

UK Oversight Pillars
Ministers
Judicial Commissioners
Intelligence and Security Committee
Investigatory Powers Tribunal
authorize surveillance warrants
audit compliance with law
oversees administration, policy, spending, activities
investigates complaints of unlawful action and human rights infringement
GCHQ

Figure 4.3. U.K. oversight pillars

4.5 What Oversight Bodies are reporting related to…

4.5.1 Effectiveness

As a whole, it was found that oversight bodies minimally treat the question of the effectiveness of surveillance programs. The PCLOB was the only oversight body found to explicitly and of its own initiative address the effectiveness of specific surveillance systems, recognizing this as a necessary step to address proportionality (PCLOB Jan. 2014 & June 2014). The Board's reviewing of these programs was initiated by requests from Congress and the President, which was a result of public outcry following the Snowden leaks. In the context of its oversight mandate related to privacy protection, the Board was tasked with reviewing two NSA surveillance programs – the bulk collection of domestic phone metadata under Section 215 of the Patriot Act, and the collection of foreign electronic communications under Section 702 of the Foreign Intelligence Surveillance Act (FISA).

The one other instance found of an oversight body evaluating the effectiveness of a surveillance system was initiated by a Defense Hotline complaint. The hotline is for reporting fraud, waste, or abuse anonymously. The complaint alleged that the NSA surveillance program TRAILBLAZER was more costly and yet inferior to the THINTHREAD system, but the NSA chose it over THINTHREAD regardless. The Inspector General of the Department of Defense consequently performed an audit on these two systems related to cost and effectiveness. Overall the audit's findings seem to agree with the complaint that TRAILBLAZER was not the best system. TRAILBLAZER was created specifically to effectively exploit the global network. It was a question of which of the two systems was most effective at achieving this goal. The report notes that a separate study "observed that the TRAILBLAZER was poorly executed" (Office of Inspector General 2004, pp.27, 29).

It is significant that this is only the second example we have of an oversight body evaluating the effectiveness of a surveillance technology, and that this evaluation was prompted by a complaint – i.e. it was not a systematic review. Likewise, the origin of the PCLOB's reports was public outcry. These effectiveness evaluations were reactive and in response to someone crying foul.

In the two reports above and in other instances where oversight bodies' reports do point to effectiveness certain trends were identified. These trends are indicated with italics. Oversight bodies place value on surveillance systems providing information that results in the *identification* of criminals and terrorists and their *plots*, *knowledge* about the functioning of their organizations, and the *prevention* of criminal acts occurring. The U.K. Surveillance Commissioner testifies in his reports of the importance of these collection programs: "I have been impressed… by how interception has contributed to a number of striking successes. It has played a key role in numerous operations including… the prevention of murders, tackling large-scale drug importations… gathering intelligence… on terrorist and various extremist organisations,… serious violent crime and terrorism" (Interception of Communications Commissioner "Report for 2006," p.11). In his 2010 report the Commissioner describes an investigation that successfully utilized interception technology, which led to members of the drug organization being identified, a better understanding of the organization's operations and interactions with other criminal organizations, the prevention of a murder, the seizure of drugs, and the arrests and convictions of principle members. The ISC judged bulk interception to be effective because it has exposed plots (ISC Mar. 2015).

Likewise, in the U.S., the Attorney General's Office argues that Section 702 is effective based on it yielding information about the identities and plans of terrorists, and the support and functioning of their organizations. It equally makes an argument for the effectiveness of the metadata collection program when it states that the FBI has opened 27 international terrorism investigations from May 2006 through the end of 2008 based, at least in part, on tips gained from this program (Assistant Attorney General 2009).

The PCLOB found the Section 215 metadata collection program to be ineffective based on, in its seven years of existence, there not being a single instance in which it significantly contributed to a counterterrorism investigation, to identifying an unknown terrorist plot, or to disrupting a terrorist attack, and there being only one instance in which a semi-unknown terrorist suspect was identified (PCLOB Jan. 2014). Conversely, the Board reported that Section 702 collection led to identifying terrorists or plots in approximately 30 cases, and that it contributed to existing investigations in about 20 cases, ultimately judging it to be "valuable and effective" (PCLOB July 2014, p.2). Additional measures of effectiveness identified are knowledge gained about the target, as well as the location and movements of suspects.

Oversight bodies also consider the number of *reports* generated to be a measure of effectiveness. That is, a surveillance program can be measured to be effective or not based on the number of reports generated containing information the system has gathered. Programs that are "effective" are implied to be those that result in reports, testimonies, and briefings of Congress (Dept. of Justice 2012). The PCLOB states that over one fourth of the NSA's reports on international terrorism "include information based in whole or in part on Section 702 collection" (July 2014, p.108). The British ISC cites the GCHQ's increased number of reports and the quality of analysis as an indication of the agency's effectiveness (ISC "Annual Report 2011-2012," p.16). This is a judgment in relation to the agency as a whole and not to surveillance technology, but it indicates what the oversight body considers to be a measure of effectiveness.

The PCLOB's report on Section 215 established seven "categories of success" by which to measure the value of a counterterrorism program (PCLOB Jan. 2014, p.146). In addition to categories covered in the preceding paragraphs these include measures of enabling *negative reporting,* adding or confirming details, and triaging. "Negative reporting" refers to establishing that a known terrorist does not have a U.S. nexus. Triaging refers to *prioritizing* leads based on urgency in a time-sensitive scenario. In the Section 702 report the Board places value on the *flexibility* of the program, which allows the government to continue monitoring suspects when they change modes of communication, and to the execution speed of the Section 702 process, which is faster than the traditional warrant process and therefore saves resources (PCLOB July 2014, pp.104-106). In reference to a certain kind of data collected, the ISC reports that this data helps the agencies quickly determine who is a potential target and who to filter out (ISC Feb. 2013). Here again value is given to *speed* and resources.

Oversight bodies rarely evaluate the effectiveness of surveillance programs themselves. In the documents analyzed, there were no indications that any of the U.K. oversight bodies explicitly perform evaluations of effectiveness. And even in the U.S. the two cases cited above – the PCLOB reports and the Department of Defense Inspector General's audit – appear to be the exceptions to how oversight bodies handle the question of the effectiveness of surveillance technology. The norm is not to perform evaluations of effectiveness themselves, but to depend on the intelligence agencies to do so.[2]

For example, the U.S. Congressional Committees do not themselves determine effectiveness, but press the intelligence community to do so. In one Senate Committee report, effectiveness is very specifically mentioned as something intelligence agencies should assess in a detailed way, to include measures of effectiveness:

> *26. Measures of effectiveness*
> *The Committee continued to press the Intelligence Community… to establish quantitative measures of effectiveness to provide insight into how effectively a program is performing… The Committee is pleased that the IC is developing more meaningful measures of effectiveness for its programs (Senate Committee Mar. 2011, p.30).*

It is worth mentioning here, a report produced by David Anderson evaluating the operational use of bulk powers used by British intelligence services (GCHQ, MI5, MI6). At the request of Parliament, Anderson reviewed these bulk powers (bulk interception, bulk acquisition, bulk equipment interference, bulk personal datasets) to assess whether they were useful for the operations in which they were used, and whether or not other techniques could have been used in their place. Anderson found these powers to be effective and necessary, using as his measure of effectiveness whether using the power in question "has made a significant contribution" to the process of identifying potential threats or sources of intelligence, understanding more fully the intelligence picture, or taking action (Anderson 2016, p.74). More specific activities identified were discovering targets, gaining knowledge about targets, detecting anomalies, analyzing networks, and triaging and prioritizing. Although not listed specifically as a measure, in his evaluation

[2] The Dutch oversight body mentioned in the introduction – CTIVD – states specifically on its website that it does not evaluate effectiveness. It is interesting that its function so definitely does not include this, and that it so pointedly wants to distance itself from any expectations of effectiveness evaluations.

Anderson clearly also places value on speed – the speed at which a given power provides the information over an alternative method. These are, obviously, many of the same measures identified above. This report is significant in that it focuses specifically on evaluating the effectiveness of surveillance programs within intelligence agencies, and it is public – a rarity – however, it is not produced by an oversight body, and thus does not strictly fall within the bounds of our research.

In the court documents analyzed for this study (i.e. those publicly available), the question of strict effectiveness is assumed by the U.S. FISC and the U.K. IPT. The documents consider the government's national security needs, but not whether the program in question is effective in contributing towards meeting those needs. The program is assumed to be effective and to therefore contribute towards national security. It is unknown whether non-public documents by the FISC might address effectiveness. In the case of the IPT it has not been called upon to address effectiveness (Anderson 2016). It is hypothetically possible that it could, although this seems unlikely since its mandate revolves around investigating complaints of unlawful action and human rights infringement.

To summarize, oversight bodies rarely treat the question of the effectiveness of surveillance technology. Rather, they expect and press the intelligence agencies to do so. When effective or successful programs are spoken of, oversight bodies place value on systems that thwart plots, provide knowledge on criminal organizations, and result in reports generated. Validating information gathered by other means, negative reporting, prioritizing leads, and speed are additional measures of effectiveness important to oversight bodies.

4.5.2 Cost

Oversight related to cost is a frequent subject in the reports of the parliamentary arms of oversight – the British Intelligence and Security Committee and the American Senate and House Committees. Their mandate includes overseeing the spending of the intelligence agencies. The ISC's annual reports review the spending of each intelligence agency. In some instances, exact amounts are classified, but the reports document the relative increases or decreases in relation to previous years. One report criticizes an agency for failing to effectively manage its expenditures for the fourth consecutive year (ISC "Annual Report 2010-2011"). Another states that the Committee's most significant concern is related to a collaborative savings program, which requires the intelligence agencies to achieve 220 million pounds of savings. It reports that "considerable improvements" are needed if the agencies are to meet this goal by the deadline (ISC

"Annual Report 2012-2013," pp.4-5). Other issues include "an SIS payment of several million pounds relating to an operation with a foreign intelligence service which was not adequately documented; spending in excess of Treasury limits on advertising and marketing" (p.34).

One of the Senate Committee's main set of reports is on the Intelligence Authorization Act for each fiscal year. While the Committee itself does not authorize spending, it reports on the authorizations and recommends whether the bill pass. Each report contains a classified section detailing the authorizations. One example of titles within the report further illustrate the cost focus: Budget and Personnel Authorizations; Increase in employee compensation and benefits; Major System Cost Reports (Senate Committee July 2010).

These parliamentary oversight bodies link the value of surveillance programs to their cost. The ISC reports on value for money and the efficiency of the agencies' spending ("Annual Report 2006-2007"). In one report it requested the National Audit Office to assess specific projects for value to money ("Annual Report 2010-2011"). In another it chides an agency for putting efficiency and "value-for-money gains at risk" ("Annual Report 2008-2009," p.16). The Senate Committee calls for vulnerability assessments of major systems in order to determine "whether funding for a particular major system should be modified or discontinued" (July 2010, p.13). Another report requests cost and feasibility studies related to the adoption of certain business systems (Nov. 13, 2013). These are assessments of the value of a program in relation to its cost.

This leads to the matter of effectiveness being discussed in the context of cost by oversight bodies. While "effectiveness" is mentioned, the focus is really more a question of cost than of strict effectiveness. The ISC reports that the government was developing "a framework for monitoring efficiency and effectiveness across the Agencies," and then goes on to discuss resources being used in an effective and efficient manner. Thus, it is actually cost-effectiveness that is being examined. The Committee assesses how the British agencies have performed, stating, "[I]t is essential that this level of funding can be justified" (ISC "Annual Report 2012-2013," p.4). Likewise, the Senate Committee, due to budget cuts, calls for data "on the effectiveness of all of the intelligence disciplines… relative to their costs to the taxpayer… Therefore, the Committee directs the ODNI to complete a detailed analysis comparing the effectiveness and costs of the Geospatial, Human, Measurement and Signatures, Open Source, and Signals Intelligence disciplines. The study must include detailed analysis of the costs and effectiveness of subcomponents and major programs" (Senate Committee Nov. 13, 2013, pp.20-21). Although effectiveness analysis is called for it is relative to and in the context of cost.

The previously mentioned Department of Defense audit of the TRAILBLAZER and THINTHREAD systems is the only example we have of an oversight body performing a cost evaluation of specific surveillance technology. The audit's findings seem to agree with the complaint that TRAILBLAZER was a costly system.[3] The report quotes another study which states "that the TRAILBLAZER was poorly executed and had an overly expensive [classified]" (Office of Inspector General 2004, p.29). The audit includes a whole classified section devoted to a cost analysis of THINTHREAD.

4.5.3 Proportionality

Ensuring that intelligence agencies stay within the bounds of the law is a central function of oversight. There are several aspects of conducting legal surveillance, such as following the correct procedure and covering only the permitted persons and communications. Surveillance can be conducted legally according to conditions such as these, and yet be disproportionate. Proportionality refers to the impact on privacy versus the benefits for security. Proportionality can be clearly built into the legal statute, or can be more vaguely referenced as something to be sought after, but not specifically required for legality. When proportionality does appear it is a sub-category of legality – as such, it is difficult to treat proportionality without also mentioning the broader subject of legality.

We observed that some statements by oversight bodies have more of a legal focus, while others concentrate on the sub-category of privacy and proportionality. Some U.S. oversight bodies seem to address either legality generally or proportionality in particular, while all the U.K. oversight bodies appear to address both broad legality and the proportionality sub-category equally.

Legality

Examples abound of various oversight bodies determining the lawfulness of intelligence agencies' actions. The ISC investigated allegations that GCHQ acted illegally in regards to accessing information gained through PRISM, and found that contrary to the allegations, GCHQ acted legally (ISC "Statement on PRISM"). The Intelligence Services Commissioner yearly reports on the lawfulness of the issuing of warrants by the intelligence agencies. In all of the IPT's judgments the Tribunal considers the legality of the actions of the intelligence agency concerned.

[3] TRAILBLAZER was eventually abandoned after over $1 billion had been spent on the program. (Mayer 2011)

The NSA Inspector General's report on the President's Surveillance Program states that the NSA General Counsel, the DOJ Office of Legal Counsel, and the NSA Inspector General all arrived at the conclusion that the President's authorization for collection of communications with one end in the U.S. was legal (NSA Inspector General 2009). The FISC found the metadata collection program to be lawful 35 times (Pauley 2013). Even U.S. Congressional Committees' jurisdiction includes aspects of legal oversight, although they do not render legal opinions (e.g. reviewing FISC orders authorizing targeted collection of communications entering or leaving the U.S. if there was probable cause of one of the parties being a terrorist (Senate Committee Mar. 2009)).

Compliance is a common subset theme of legality. Certain oversight bodies produce yearly compliance reports. For example, the Attorney General and DNI jointly produce a semiannual assessment of the NSA's compliance with procedures and guidelines related to Section 702. These reports detail and number the errors, and determine whether or not intentional violations have been made. Based on one declassified report we can deduce that these reports also detail the more significant incidents of non-compliance, i.e. those involving U.S. persons (Attorney General and DNI 2013).

The U.K. Interceptions of Communications and Intelligence Services Commissioners report on compliance that relates to their respective jurisdictions. Both produce an annual report documenting the number of errors made by the intelligence agencies. They contextualize the errors by categorizing them and highlighting the severity of the error and the degree of privacy intrusion; examples of errors are also detailed. One such example is that of a GCHQ internal monitoring system of staff communications capturing more information than authorized. The Commissioner concluded this was a technical error. GCHQ deleted the relevant data and reconfigured the system to ensure compliance (Intelligence Services Commissioner 2015).

U.K. and U.S. reports are similar in that they both discuss the types of non-compliance, the number of incidents separated by agency, give examples of the errors, and describe what action was taken to correct the error and to prevent it from happening again. The difference is that the U.K. reports are originally intended for the public, while the U.S. reports are classified.

The term "compliance" is used in discussing errors made that subsequently mean that the agency's actions were not according to the law. Interestingly, it seems that "legal" is used most often to refer to a surveillance program as a whole or the carrying out of surveillance duties as a whole. "Compliance" is used primarily to refer to the mistakes made within these legal programs.

The oversight documents that report on compliance also address *integrity* and whether or not the errors were intentional. In all the reports reviewed, the oversight mechanisms stressed the integrity of the intelligence personnel and their desire to act within the law. The language was found to be slightly stronger when the report's audience was the public at large.

The Interception of Communications Commissioner reported that he found no evidence of a desire to act unlawfully within the intelligence agencies, but rather a clear desire to ensure that their actions are within the law (2007). The Intelligence Services Commissioner takes care to stress that in instances of non-compliance, "None of the cases involved bad faith or any deliberate departure from established practices" (July 2008, p.9). The Attorney General and DNI state that NSA agency personnel demonstrate "a focused and concerted effort" to comply with requirements, and report that they found no intentional violations in the instances of non-compliance (2013, pp.22, 37).

Privacy and proportionality

Proportionality falls within this broader theme of legality. It is closely associated with the notion of privacy protection, the (widely-held) belief being that if surveillance is proportional, the privacy of innocent citizens will better be protected. In the U.S. it is most often the PCLOB and FISC that address proportionality and privacy. In both the U.S and the U.K., oversight bodies seek for *privacy protections* to be built into surveillance systems. The ISC called for privacy protections to "form the backbone" of new legislation being drafted for investigatory powers, and not be handled as a mere "add-on" (ISC Feb. 2016, p.3). The Interceptions of Communications Commissioner considers it his role to ensure that systems are in place to protect the privacy of British citizens (2007).

The FISC imposes certain measures on NSA surveillance systems to protect privacy, such as instituting regulations regarding accessing and storage of metadata, as well as requiring random spot checks and authorizations for certain activities. The Court recognizes that the data collected by the NSA under the telephone metadata program will be broad, but qualifies that "the use of that information for analysis shall be strictly tailored to identifying terrorist communications" and must be carried out according to prescribed procedures (Judge FISC May 3, 2007, p.7). The NSA is only permitted to search this metadata when it has reasonable suspicion that a telephone number is associated with a terrorist suspect.

While privacy controls are in place for various surveillance systems, our research revealed that whether these controls are adequate, and whether in given instances the privacy invasion is proportionate to the security concern, is a matter of *human judgment*.

This judgment is passed by judges, commissioners, oversight committees, and board members.

The ISC judges that the privacy concerns of examining datasets containing large volumes of data of people of non-interest outweigh the practical considerations (significantly increasing the number of warrants and therefore also time and cost) of the intelligence agencies; the intrusion merits requiring a specific warrant (ISC Feb. 2016). The Intelligence Services Commissioner testifies that the question he focuses on in his oversight is that of proportionality, assessing whether the agencies have correctly balanced the security necessity against the privacy invasion (ISC Dec. 21, 2015, p.806).

In a report to the FISC, the Attorney General makes a proportionality judgment supporting the NSA's balancing of security needs, cost, and protection of privacy: destroying credit card information contained in call records requires personnel, time and resources, which "are not justified given the operational need for certain information" and the measures taken to ensure the records are secure (Assistant Attorney General 2009, p.65). A FISC judge finds that a two-year rather than five-year retention of upstream acquisitions "strikes a more reasonable balance" between security needs and protecting privacy (Judge FISC Oct. 2011, p.13).

The PCLOB found the Section 215 program to be disproportionate – one instance of identifying a not entirely unknown terrorist suspect hardly justifies the broad collection of phone metadata. This instance, in particular, is a good example of human judgment at play, because two of the Board's five members wrote dissenting opinions disagreeing with the conclusion that Section 215 was disproportionate.[4] One found that the limited amount of information collected by the program, along with the existing and PCLOB-recommended privacy protections renders the privacy intrusion small, while the potential benefit of the program remains significant (PCLOB Jan. 2014).

All the above examples indicated that while laws and measures may be in place to protect privacy and ensure proportionality, the actual determination of whether or not proportionality is achieved must ultimately fall to individual human judgment. And, naturally, different individuals will often arrive at different conclusions.

Proportionate ↔ legal

One finding of this study is that there is a significant interplay between legality and proportionality. In the U.K. case, what is legal is determined, in part, by what is

[4] They also disagreed with the conclusion that Section 215 is not an effective program.

proportional. While proportionality is not the only aspect that determines legality, without proportionality being achieved neither can legality be achieved. U.K. law stipulates that any surveillance performed with surveillance technology must be shown to be firstly necessary, and secondly proportionate to what it seeks to achieve. Thus, the case for proportionality is built directly into the law. The IPT states that "indiscriminate trawling for information… would be unlawful" (IPT June 2015, p.77). "Indiscriminate trawling" is considered to be disproportionate which therefore makes it unlawful.

Across the Atlantic, the role of proportionality is more vague. The FISC states that to assess reasonableness (proportionality), a court must consider "the nature of the government intrusion and how the government intrusion is implemented. The more important the government's interest, the greater the intrusion that may be constitutionally tolerated" (Judge FISC Oct. 3, 2011, pp.69-70). The court goes on to state that if the privacy protections are adequate "the constitutional scales will tilt in favor of upholding the government's actions" (p.70). If, on the other hand, the protections are inadequate to protect against the risk of error and abuse, the balance will tip towards a judgment of unconstitutionality. Determining whether a program is constitutional or legal, therefore, includes determinations of proportionality. The law requires U.S. intelligence agencies to implement measures to protect privacy. Whether or not a given program is found to be legal or not, however, is based upon whether these measures are deemed to be adequate in light of the government's interests. If a program is determined to be proportional it is considered constitutional. If it is determined to be overly invasive it will be found to be unconstitutional.

4.6 Discussion

This section contains themes that were identified across all three elements of effectiveness, cost, and proportionality, and compares the results of this paper with the authors' previous study on intelligence practitioners.

While this study intentionally covered a pre- and post- Snowden timespan, no significant differences were found in how oversight bodies dealt with the three aspects of overall effectiveness (strict effectiveness, cost, proportionality) in these two different periods.

4.6.1 Oversight bodies and intelligence practitioners compared

Comparing the results of this study with the findings regarding intelligence practitioners revealed some notable similarities and differences.

The question of the effectiveness of surveillance technology is rarely treated by oversight bodies and intelligence practitioners alike. In their investigation of two of the NSA's surveillance programs, the PCLOB specifically raised the question of whether these programs were effective. The Board's hearings, in which they posed this question to intelligence officials, is the sole instance found of intelligence practitioners specifically addressing this question. The PCLOB's reports are also only one of two instances found of oversight bodies doing so. Instead, discussions turn around cost or privacy and proportionality issues, or focus on existing or new oversight mechanisms to implement.

Among the measures of effectiveness identified, oversight bodies and intelligence practitioners signal several of the same measures. Thwarting plots, identifying and locating criminal and terrorist suspects, and providing knowledge of the structure and workings of criminal organizations are considered by both to be ways of evaluating effectiveness. The number and quality of reports generated based on information gained from a surveillance program is also considered to be an important measure of effectiveness by both stakeholders.

Both groups also revealed cost as a driver of evaluations of effectiveness. Effectiveness is evaluated not so much out of concern for effectiveness itself, but out of concern for cost. Oversight bodies call for assessments evaluating value to cost, and practitioners evaluate systems because both parties only want money spent on programs that are effective.

Proportionality judgments involving individual human judgment was a theme apparent in both studies. It is individual human beings who ultimately decide what is proportional both within oversight mechanisms – courts, boards, legal offices – and within the intelligence agencies – e.g. directors.

One notable difference is in the interplay between legality and proportionality. Intelligence practitioners state that proportionality is determined by the law: they themselves do not make proportionality judgments; they simply act according to what the law prescribes. The law determines what is proportional, and this is enforced by oversight. Our research findings on oversight bodies seem to reveal the reverse of this logic. That is, that legality is determined, in part, by what is proportionate. The law lays down certain procedures (e.g. FISC guidelines), but there is still room for, and it is even necessary to have, judgments of what is proportional. These proportionality judgments are part of what determine legality.

4.6.2 Dependency

Although many oversight elements are independent of intelligence agencies and therefore their work is independent, our research revealed that they are heavily dependent on the intelligence community for the documents and testimony necessary to carry out their oversight. Likewise, they are dependent on intelligence agencies to determine the effectiveness of surveillance technology. The technical expertise necessary to perform these evaluations lies largely within the intelligence agencies. While some oversight bodies draw on outside technical support (e.g. the PCLOB held a public forum which included a panel of technology experts), these experts do not have access to classified material and therefore are not advising specifically on the surveillance systems in question.[5] A notable exception is the IOCCO, which includes technical experts as part of its inspection team. However, the IOCCO does not explicitly evaluate effectiveness.

The Senate Committee's conducting of its oversight of NSA electronic surveillance was assisted by briefings by the NSA and access to court documents (Senate Committee Mar. 2009). The PCLOB's reports on NSA surveillance programs relied on testimony, hearings, and evidence received from the members of the intelligence community. Likewise, the ISC concluded that media allegations that GCHQ circumvented British law were false based on evidence given by GCHQ (ISC "Statement on PRISM").

This inter-relatedness is also evident in the measures of effectiveness unearthed in our two studies. At least part of the reason both practitioners and oversight bodies come up with similar measures of effectiveness, is that the oversight bodies are relying on intelligence officials to indicate how to evaluate the effectiveness of surveillance programs.[6]

This dependency does not equate to oversight bodies giving the agencies a green pass at every turn. Examples abound of oversight bodies finding fault with surveillance programs: the audit on TRAILBLAZER, the PCLOB 215 report, FISC and IPT judgments, and compliance issues raised by the FISC and the U.K. Commissioners. It is, however, an interesting point that, in order to conduct their oversight (including any conclusions on effectiveness), oversight bodies must rely on the intelligence agencies themselves for testimony, documentation, error reporting, and the like. This dependency seems inevitable. It is the members of the intelligence agencies who carry out the surveillance

[5] An exception to this was Anderson's independent review, whose team, which had access to classified documents, included a technical expert. However, as stated previously, this review was not performed by an oversight body.

[6] This is also true of Anderson's independent review.

actions and use the technology. Arguably, therefore, they know best the functioning of the systems, what actions they have taken, errors they have made, and the subsequent documentation. This inter-reliance, however, explains why certain groups and individuals claim that oversight bodies no longer serve their purpose and have been co-opted (e.g. Greenwald). It also points to an issue of trust.

4.6.3 Trilemma

Many oversight bodies are given a mandate that focuses on one of the three elements of effectiveness, cost, and legality/proportionality.[7] For example, the IPT was established to handle complaints regarding unlawful or disproportionate actions by the intelligence agencies. The NSA General Counsel is charged with providing legal advice. Consequently, their activity focuses on the given element. Any given report tackles only one and occasionally two of these elements together: the U.K. Commissioners' reports address compliance; the ISC annually reviews the intelligence agencies' spending; the PCLOB reports on NSA surveillance programs analyze effectiveness and legality; the Inspector General's audit on TRAILBLAZER and THINTHREAD reviews the cost and effectiveness of the programs. The Bulk Powers Review, although not produced by an oversight body, was launched to investigate effectiveness and specifically excluded proportionality. A final example is that of the Senate Committee initiating an in-depth review of the legality and cost-effectiveness of intelligence collection programs (Senate Committee 2015). This example is particularly interestingly because it is a review specifically of surveillance programs and it focuses on cost and legality, but not on strict effectiveness. In the documents studied, oversight bodies were never found to address all three elements of effectiveness, cost, and proportionality simultaneously. Mechanisms designed to deal with all three, like the Senate and House committees, never evaluated all three at once.

The fact that no agency or oversight body addresses all three elements together reminds us of a well-established theory in macroeconomics – the impossible trinity, or trilemma. This theory states that policymakers in open economies must choose two out of three conflicting, yet desirable goals: monetary independence, exchange rate stability, and financial integration. Because it is impossible to have all three, policymakers must decide which one they will give up (Aizenman and Ito 2012, Obstfeld et al. 2004). No such formal framework exists in the security realm; however, the same reality is present. While oversight bodies (and intelligence practitioners) speak of simultaneously

[7] Continuing from section 4.5.3 on proportionality, which covered the categories of *legal* and *compliance*, the proportionality element here includes legal and compliance.

delivering effective surveillance, in a cost-efficient manner, while maintaining proportionality, these are, in fact, conflicting goals.

Both domains contain three conflicting goals, and in practice, stakeholders address only two of the three goals simultaneously. This trilemma concept points to why many oversight bodies are tasked with performing only one of these activities, such as overseeing the protection of privacy or of legal compliance. Others are tasked with two or all three missions, but alternately perform them one (or possibly two) at a time. This allows oversight mechanisms to successfully treat the issue at hand without having to enter into the impossible task of successfully addressing all three elements.

4.7 Conclusion

As digital data has become increasingly important to society, so too has it become central to the work of intelligence agencies. As their surveillance of this data increases, so does the importance of the work of oversight bodies. Consequently, how oversight bodies consider and evaluate the overall effectiveness (including effectiveness, cost, and proportionality) of surveillance technology is a crucial question.

Oversight bodies were found to minimally treat the question of strict effectiveness. Instead they rely on the intelligence agencies to perform evaluations of effectiveness. Measures of effectiveness that oversight bodies were found to value are: thwarted plots, knowledge gained, reports, validating and prioritizing information, and speed. These are similar to the measures identified for intelligence practitioners, pointing to a dependency of oversight bodies on intelligence officials to indicate how to evaluate surveillance programs. Oversight bodies are equally dependent on the agencies for the documentation and testimony necessary to perform their oversight.

Overseeing spending is the specific mandate of certain oversight bodies. In this context they speak of the value of surveillance programs in relation to their cost. Evaluations of surveillance technology focus on cost to value considerations. This is another similarity found with intelligence practitioners.

Ensuring agencies and their surveillance technology stay within the bounds of the law is an important and central function of oversight. In addition to ensuring the legality of surveillance programs, oversight mechanisms report on compliance, enumerating and investigating errors made within legal programs. Oversight bodies ultimately rely on judgments of proportionality to help determine legality, while intelligence practitioners demonstrate the reverse, relying on the law to determine proportionality.

Oversight mechanisms typically have a mandate concerned with one of the three elements of effectiveness, cost, and proportionality. If their mandate includes more than one of these three, in any given report they only evaluate one or two of the three, but never all three simultaneously. Successfully addressing all three is an impossible trilemma.

The results of this study, along with those of the authors' previous paper, are an important component of the ongoing discussion surrounding surveillance technology. Understanding how intelligence practitioners and oversight bodies treat questions of effectiveness, cost, and proportionality, and weigh these elements against one another, is crucial to creating meaningful dialogue between these groups and others, such as the public and privacy advocate groups. Such dialogue is necessary to build trust, without which the use of surveillance technology might undermine the democratic culture it is meant to protect.

Chapter 5
The Public Wants It All[1]

5.1 Introduction

Amidst the anguish and anger following terrorist attacks such as 9/11 and Charlie Hebdo, erupt calls for more government action to prevent such events. This, in turn, leads to increased surveillance by intelligence agencies. Months or years later, leaked classified documents, such as in the Snowden case, result in privacy advocates decrying the government's surveillance actions as a violation of privacy. A debate ensues in which security is pitted against privacy. Such is the current discussion on government surveillance, which turns around privacy versus security.

The public is one actor in this debate. On both sides, other actors claim to act in the public's interest, either accusing the government of privacy rights' violation and demanding a cessation of activities, or defending these surveillance programs as necessary to keep citizens secure against current threats. Both sides seek to influence whether or not the public accepts these surveillance programs, basing their arguments in the context of the privacy vs. security trade-off.

Within the privacy-security debate there is a less-heard debate: that of the effectiveness of the surveillance programs in question. Is the surveillance technology effective in accomplishing a given security goal? Different actors hold different views, and how to evaluate effectiveness is in itself a complex question (Cayford and Pieters 2018). The public's perception of effectiveness has significant social implications, as this potentially plays a role in its acceptance of the technology; judgments on acceptability may depend on judgments of effectiveness. This, in turn, could influence the legitimacy and continued use of such technology. In democratic societies the public's view matters. Its views influence policy and the making of law.

This study explores how the public perceives the effectiveness of surveillance technology, and how people's views on privacy and their views on effectiveness are related. Likewise, is there a relation between perceptions of effectiveness and opinions

[1] This chapter was published as Cayford, Michelle, Wolter Pieters, and P.H.A.J.M van Gelder. "Wanting it all – public perceptions of the effectiveness, cost, and privacy of surveillance technology," *Journal of Information, Communication and Ethics in Society*, 18:1, 10-27. 2019. DOI: 10.1108/JICES-11-2018-0087.

on acceptable cost of surveillance technology? This paper analyzes the results of surveys completed by a group of Dutch undergraduate students and their parents. It is one of only two studies investigating views of the Dutch public in relation to surveillance, and the first known study to analyze public views regarding effectiveness and the correlations between effectiveness, privacy, and cost. Findings show some significant differences between the parent and student generations, as well as correlations that suggest that the public does not engage in the trade-offs imbedded in the privacy-security debate. Rather, the public wants it all – effectiveness and privacy at a reasonable cost, all delivered simultaneously.

The paper is structured as follows: Section 5.2 presents related work, followed by methodology in Section 5.3. The study's results are then given accompanied by analysis in Section 5.4, followed by a discussion of the results in Section 5.5 and finally, conclusions in Section 5.6.

5.2 Related Work

The question of the effectiveness of surveillance technology is closely linked to the debate over the values of security and privacy. On one side of this debate are those who argue that access to telecommunications is necessary for stopping organized crime and terrorism. On the other side of the debate are those who argue for the right to privacy. The privacy-security debate is expansive, and the literature reflects this, spanning discussions of privacy and security and technology (Friedewald and Pohoryles 2013, Schneier 2013), privacy and data retention law (Tene 2011, Mitrou 2007), the question of balance that should be struck between the two (Guerrier 2016, Stalla-Bourdillon et al. 2014), whether or not there is or should be a trade-off between the two (Verfaillie and van den Herrewegen 2013, Van Lieshout et al. 2013), ethics (Stahl 2007, Landau 2014), dealing with cryptology (Diffie and Landau 2007), and the balance between liberty and security with regard to law (De Hert 2005, Poullet 2004).

This debate enters the political sphere as countries decide which balance to strike and accordingly adopt laws that restrict or allow, as the case may be, intelligence agencies' access to communications (Mansfield-Devine 2015). This is where the people come in – in democratic societies the people's view of intelligence agencies and the surveillance technology they employ affects this debate. If the public perceives surveillance technologies as effective and the associated agencies as trustworthy, negative impacts on privacy are more likely to be accepted.

Public perception is a topic in and of itself, which includes actual versus perceived effectiveness, and influences on perceptions such as politics, the media (Altheide 2006), culture, place and situations (Orru 2013). Here, the goal is not to focus on these influences, nor to enter into a discussion of perceived versus actual effectiveness. Rather, it is to continue previous research of investigating different stakeholders' views of effectiveness. All stakeholders perceive effectiveness in some way, and these various views influence the debate around the use of surveillance technology.

A review of the existing literature treating surveys of the public related to surveillance reveals that they primarily focus on privacy issues, and that many are concentrated around the Snowden leaks. One such study surveyed students in various countries (Germany, Spain, Sweden, Japan, China, Taiwan, Mexico, and New Zealand) regarding their views on privacy and state surveillance following the Snowden leaks (Adams et al. 2017, Murata et al. 2017, Kavathatzopoulos et al. 2017, Gunasekara et al. 2017). The study's questions probed general privacy attitudes and investigated knowledge and evaluation of Snowden's actions. The surveys found, among other things, that attitudes toward privacy varied in different countries, and that most respondents approved of Snowden's actions and a majority would act as he had, except for those in China and Japan. A separate study surveyed students in Slovenia in order to understand public perception of cyber-surveillance pre- and post-Snowden, and found that more people felt threatened by domestic and foreign intelligence services post-Snowden (Zavrsnik and Levicnik 2015).

The Pew Research Center in the U.S. surveys public opinion and has produced several reports of survey results related to National Security Agency (NSA) surveillance, privacy, and Snowden. One such report revealed that a small majority (50%) approved of the government surveillance programs collecting phone and internet data (Pew July 26, 2013). Another survey (Madden 2014) indicated that Americans are concerned about government and businesses' surveillance of their communications. A Pew report focused on Americans' views of NSA surveillance and privacy and security (Gao 2015) found that a majority do not believe they need to sacrifice privacy and freedom to be safe from terrorism, while also stating that anti-terrorism policies do not do enough to protect them. Additional Pew surveys polled Americans' views on NSA surveillance programs and of Snowden's actions (Desilver 2014, Pew June 10 and 17, 2013), of monitoring Allied leaders' phones (Pew Nov. 4, 2013), whether they think the government is monitoring their calls and emails (Olmstead 2017), and their concerns of security versus privacy (Doherty 2013, Rainie and Maniam 2016, Rainie 2016).

Outside of student surveys and Pew surveys, Reddick et al. (2015) examine American public opinion related to NSA surveillance programs, studying the correlation between engagement in political discourse (and therefore political efficacy) and approval of these programs. Brooks and Manza (2013) write about American public opinion on counterterrorism. Through surveys conducted in 2007, 2009, and 2010 covering ten counterterrorism policies and practices, their research found strong support for NSA surveillance of telephone conversations between U.S. citizens and suspected terrorists, as well as for the Patriot Act, which was described as facilitating government access to phone and email records. The authors believe that the persistence of the high-intensity counterterrorism strategies put in place following 9/11 are in large part explained by high levels of public support.

In all the above studies and surveys, themes emerge of investigating views on privacy, acceptance of surveillance programs, and the impact of the Snowden leaks. They tend to be conducted within the context of and with an implicit understanding that citizens weigh and exchange their privacy vs. their security. This is the notion of trade-off. This notion implies that in order to have more privacy some security must be given up and vice versa. A small set of studies probes this understanding, using surveys to examine how citizens assess surveillance technology, including whether they evaluate surveillance technologies according to the security-privacy trade-off model.

The first study (Pavone and Degli Esposti 2012), suggests that the public's assessments are context-specific, reflecting their trust or distrust of the institution conducting the surveillance, and that the public does not perform a trade-off. Those who are distrusting see their privacy being infringed upon by the surveillance technology without their security being increased – in other words, the technology is both ineffective and privacy-invasive – while those who are trusting considered the technology to be effective without infringing on their privacy. Through a citizen summit in Spain, Degli Eposti and Santiago Gomez (2015) investigated the public's perceptions of CCTV and deep packet inspection (DPI), including whether they considered these two technologies effective national security tools. The study found that more participants (63%) considered CCTV to be an effective security tool than DPI (43%), and that half the participants held a security-privacy trade-off approach, considering these technologies to both enhance security and invade their privacy. The rest either viewed the technology as highly invasive and not really effective, or as very effective without infringing on privacy. Finally, Degli Esposti et al. (2017) report on surveys conducted in six European countries investigating public perceptions of DPI. They found that a technology's perceived intrusiveness negatively impacts its perceived effectiveness, and that security and privacy are "compatible rather than antagonistic dimensions" (p.72).

Similar to the above studies, Van den Broek et al. (2017) surveyed EU citizens using real-life scenarios of security issues, but without explicit references to security or privacy. Among the authors' conclusions were that acceptance of surveillance and views of privacy intrusion depend on the security issue, that trust plays a role, and that those with a high degree of privacy awareness are less likely to be accepting of surveillance while those with a high level of security concerns are more accepting. The authors conclude that "EU citizens assess security and privacy aspects as rather independent values that both need to be secured" (p.29).

This current research contributes to the work investigating the public's views on surveillance technology, zeroing in specifically on the question of effectiveness, and the relationship between perceptions of effectiveness and privacy. As with the notion of the privacy-security trade-off, it examines if the public engages in a privacy-effectiveness trade-off, exchanging more effectiveness for less privacy and vice versa. This study also investigates how the public views the cost of surveillance technology and if there is an effectiveness-cost relation in which the public is willing to spend more money if it considers the surveillance technology to be effective and vice versa.

In addition to governments, internet service providers, businesses, employers, advertising, social media, and personal devices all introduce and perform forms of surveillance (Asghari et al. 2012, Bendrath and Mueller 2011, Trottier & Lyon 2012, Trottier 2012, Fuchs and Trottier 2017). This study also explores whether people differentiate between types of surveillance – whether they feel differently about privacy and effectiveness with government surveillance versus commercial surveillance versus individuals' surveillance using mobile devices.

5.3 Methodology

This paper analyzes the results of a survey, designed to explore the public's views on surveillance technology in relation to effectiveness, privacy, and cost. The survey polled a group of bachelor's students and their parents in 2014, 2015, and 2016. The students were part of a Safety, Security, and Justice (SSJ) minor run jointly by Delft University of Technology and Leiden University in the Netherlands. In 2014, SSJ minor students were enlisted to distribute surveys to non-SSJ minor students and their parents. Both 43 non-SSJ minor students and 43 of their parents completed the survey. In 2015, only students completed the survey. In total, over the three years, there were 359 respondents, of which 200 were students. Among the respondents, 164 were females and 195 were males. The survey was conducted in English.

The survey's first three questions explored how bothered people were by different parties – ISPs; the Dutch intelligence services, Algemene Inlichtingen- en Veiligheidsdienst (AIVD); and the American NSA – potentially having access to their online activity. The aim was to explore differences and relations between how people felt about privacy intrusion by private companies versus domestic intelligence versus foreign intelligence services.

Two yes/no questions explored attitudes related to the effectiveness of surveillance technology. In order to learn more about why participants considered a particular NSA technology to be effective or not, we added a free text response. Three additional questions focused on privacy attitudes related to individuals' use of their mobile devices to record and report crime and speeding. Likert scale responses were used to gauge the degree to which respondents were bothered or comfortable with different surveillance actions.

Two years into conducting the survey, our research on the effectiveness of surveillance technology was indicating cost to be an important question with other stakeholders. A question on cost was, therefore, added in 2016 to explore attitudes of the public related to the cost of surveillance technology. The final two questions investigated expectations of security, with the intention of studying any correlations with this and perceptions of effectiveness.

As our data were not normally distributed, non-parametrical statistical tests were used to test the above hypotheses. Mann-Whitney tests were used to compare means between different groups of respondents. Spearman tests were used for correlations. Wilcoxon signed ranks test and sign test were used for pairwise comparisons. Z-tests were used for comparing population proportions, and a parametric Levene's test was used for comparing the variance between different subpopulations. The null hypothesis is rejected if the *p*-value (probability of rejecting the null hypothesis given that it is true) is less than a predetermined significance level of 5%. Exact *p*-values will be reported, unless the *p*-value is below 1‰[2]. In that case the *p*-value will be reported as p<.001. Finally, to analyze the free-text responses, the responses were coded to identify themes.

5.4 Results and analysis

This section presents and analyzes the results, with the associated hypotheses presented at the beginning of each sub-section.

[2] Due to an editing error, this was incorrectly reported as "1%" in the published version of this chapter.

5.4.1 Privacy

It was expected that attitudes regarding privacy would vary depending on who was accessing people's data. This is inherently linked to a question of trust – those who trust a given institution are more likely to be less bothered by potential privacy intrusion.

H1. *People are more bothered by the NSA potentially accessing their online activity, than by the AIVD.*

H2. *People are more bothered by both the NSA and the AIVD potentially accessing their online activity, than by ISPs and internet companies.*

H3. *Parents, being closer to a pre-internet era and a time when intelligence agencies were more trusted, will be less bothered than students by potential NSA and AIVD access, and more bothered than students by ISP and internet companies' access.*

The response results of the first three survey questions are shown in Figure 5.1. From left to right, the bars show respondents being least bothered ("not at all," corresponding to a score 1) to most bothered ("so much that I would or have taken action," corresponding to a score 4). For Q1 (*M*= 2.20, *SD* = .797) most respondents reported being "a little" bothered by ISPs and internet companies having access to their online information – 47%. Q2 (*M*= 2.15, *SD* = .916) also showed the highest number of respondents being "a little" bothered by AIVD access – 37%. While the means for Q1 and Q2 were similar, the figure suggests that responses for the AIVD are more spread. Indeed, the standard deviation for Q1 was .797, and that for Q2 was .916, meaning that responses for Q2 tended more toward the extremes, with more people reporting that they were "not at all" bothered by the AIVD and that they were so bothered that they had taken action, such as using encryption, than those who reported the same for ISPs and internet companies. In regards to the NSA Q3 (*M*= 2.54, *SD* = .918) most were bothered "a lot" – 42%. As expected, respondents were significantly more bothered by the NSA than by both private companies and the AIVD ($p < .001$), according to a pairwise Wilcoxon signed rank test.

Mann-Whitney tests were run to test hypothesis *H3*. While there was no significant difference in the means between students and parents for Q1 and Q2, the standard deviation showed differences in variance. Related to internet providers and companies, the 197 students (*M* = 2.18) had a standard deviation of .726, while the 160 parents (*M* = 2.21) showed a higher standard deviation of .879 ($p < 0.001$). As regards the AIVD, 199 students showed, *M* = 2.15, *SD* = .851, and 160 parents also showed a higher standard deviation, *M* = 2.17, *SD* = .992 ($p < 0.001$).

One of the goals of this study was to investigate relationships between responses related to effectiveness, privacy, and cost. This included correlations in responses to questions on the same subject, such as privacy (hypotheses *H1-H3*). This was done by using Spearman's correlation coefficients. These tests did indeed reveal some interesting correlations. There was a significant correlation between Q1 (ISPs) and Q2 (AIVD) (*r(357)* = .467, $p < .001$), as well as between Q1 and Q3 (NSA) (*r(357)* = .460, $p < .001$). The correlation of how people answered Q2 and Q3 was even stronger, *r(358)* = .624, $p < .001$. This suggests that those who are bothered by others having access to their online information tend to be bothered regardless of whether it is an ISP or internet company, a domestic government, or a foreign government. And likewise, those that are more accepting, are accepting regardless of the party accessing the data.

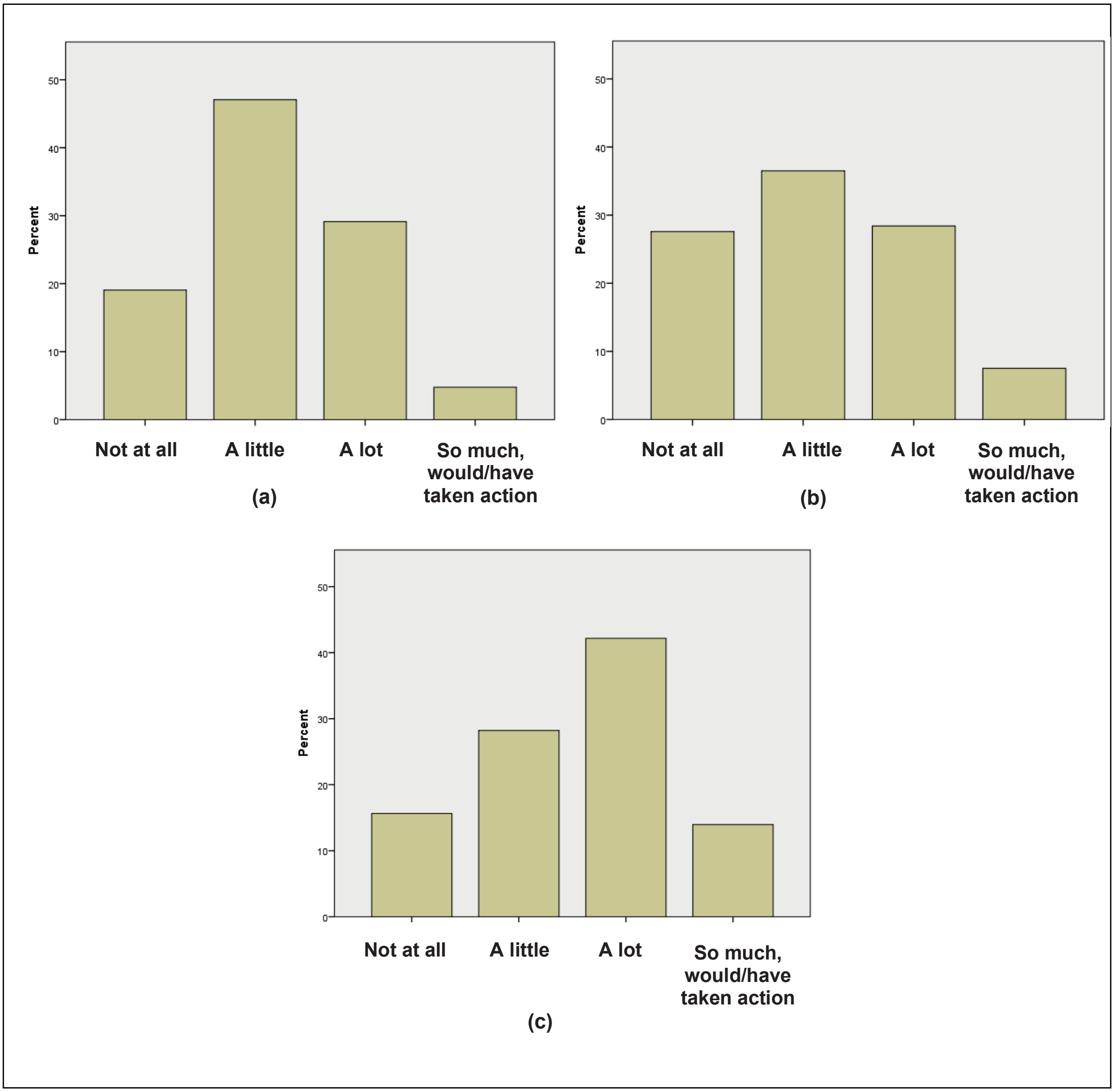


Figure 5.1.

(a)(Q1) To what extent are you bothered by Internet Service Providers and Internet companies having access to all your online activity?

(b) (Q2) To what extent would you be bothered if the Dutch AIVD had access to all your online activity?

(c) (Q3) To what extent would you be bothered if the American NSA had access to all your online activity?

5.4.2 Effectiveness

This study anticipated that attitudes concerning privacy invasion would correspond with perceptions of effectiveness of the institution or technology involved.

H4. *Those who find NSA surveillance to be effective are less bothered by potential NSA access to their online activity.*

H5. *Those who find NSA surveillance to be effective are less bothered by potential AIVD access to their online activity.*

H6. *Those who consider using private mobile devices to record crime scenes and speeding vehicles to be effective for security find this action to not be particularly invasive.*

H7. *Those who consider using private mobile devices to be effective for security are more comfortable with the police using such recordings to issue speeding fines and investigate crime.*

The two survey questions specifically addressing the question of effectiveness produced fairly different results [Figure 5.2, graphs (a) & (b)]. Respondents were nearly equally divided regarding whether NSA wiretaps with wide filters were an effective form of surveillance technology, with 50% responding “yes,” and 47% responding “no” (3% gave no answer). However, many more thought that use of mobile devices by individuals to record crime and infractions was an effective use of technology for security purposes. Close to 80% responded “yes,” with a bit over 20% saying “no.” Most people found this to be less or as equally invasive as CCTV (*M*= 1.83, *SD* = .740). Only 20% reported it was more invasive [Figure 5.2, graph (c)]. There was considerable difference, however, between how comfortable people were with police using videos from private individuals to issue speeding fines versus investigating crime [Figure 5.2, graphs (d) & (e)]. A comparison of means showed a strong significant difference with $p < .001$. Only 14% were “very comfortable” and 26% “a little comfortable” with police using private videos to issue speeding fines. The rest were closely divided between using such videos “only in cases of extreme speeding,” or “not at all.” In contrast, regarding police using private videos to investigate crime, a strong majority was “very comfortable” at 44%, and 34% were “a little comfortable.” The remainder were equally split at 11% each, with being comfortable only in cases of bodily harm or not at all. A difference in the spread of responses between students and parents emerged related to police using private videos to issue speeding fines with parents showing more variance than students ($p = .045$): 198 students (*M* = 2.78, *SD* = .998) and 160 parents (*M* = 2.76, *SD* = 1.103). This same difference in variance did not emerge related to police using private videos to investigate crime.

It was anticipated that those who considered NSA wiretapping surveillance technology to be effective would also be less bothered by NSA access to online activity (hypothesis *H4*). Indeed, there was a significant correlation between the responses, although this correlation appears to be associated with age/generation. Parents showed a strong correlation, *r(151)* = .278, *p* < .001, while the student population had no significant correlation. Although the question on effectiveness was specifically related to the NSA, we hypothesized that there would still be a correlation between this question and that dealing with privacy related to the domestic intelligence agency (hypothesis *H5*). This was, in fact, the case, with those who consider wiretaps to be non-effective, also being more likely to be bothered by AIVD access to their online activity (*r(348)* = .120, *p* = .026). Once again, this correlation is associated with age – when looked at separately, students had no significant correlation, while their parents showed a strong correlation, *r(151)* = .229, *p* = .005. Although there was not expected to be a correlation between views on the effectiveness of NSA surveillance technology and ISP access to online activity, this same pattern was found of there being no significant correlation with students, but a significant correlation existing with their parents, *r(151)* = .165, *p* = .043.

The study's second question on effectiveness concerned the effectiveness of recordings on private mobile devices for security purposes. As expected (hypothesis *H6*), those who consider private recordings of crime to be effective for security, found this action to not be particularly invasive – less invasive than CCTV (*r(358)* = .124, *p* = .019). Likewise (hypothesis *H7*), those who considered it effective were more likely to be more comfortable with police using private videos to issue speeding fines (*r(357)* = .205, *p* < .001) and investigate crime (*r(356)* = .226, *p* < .001).

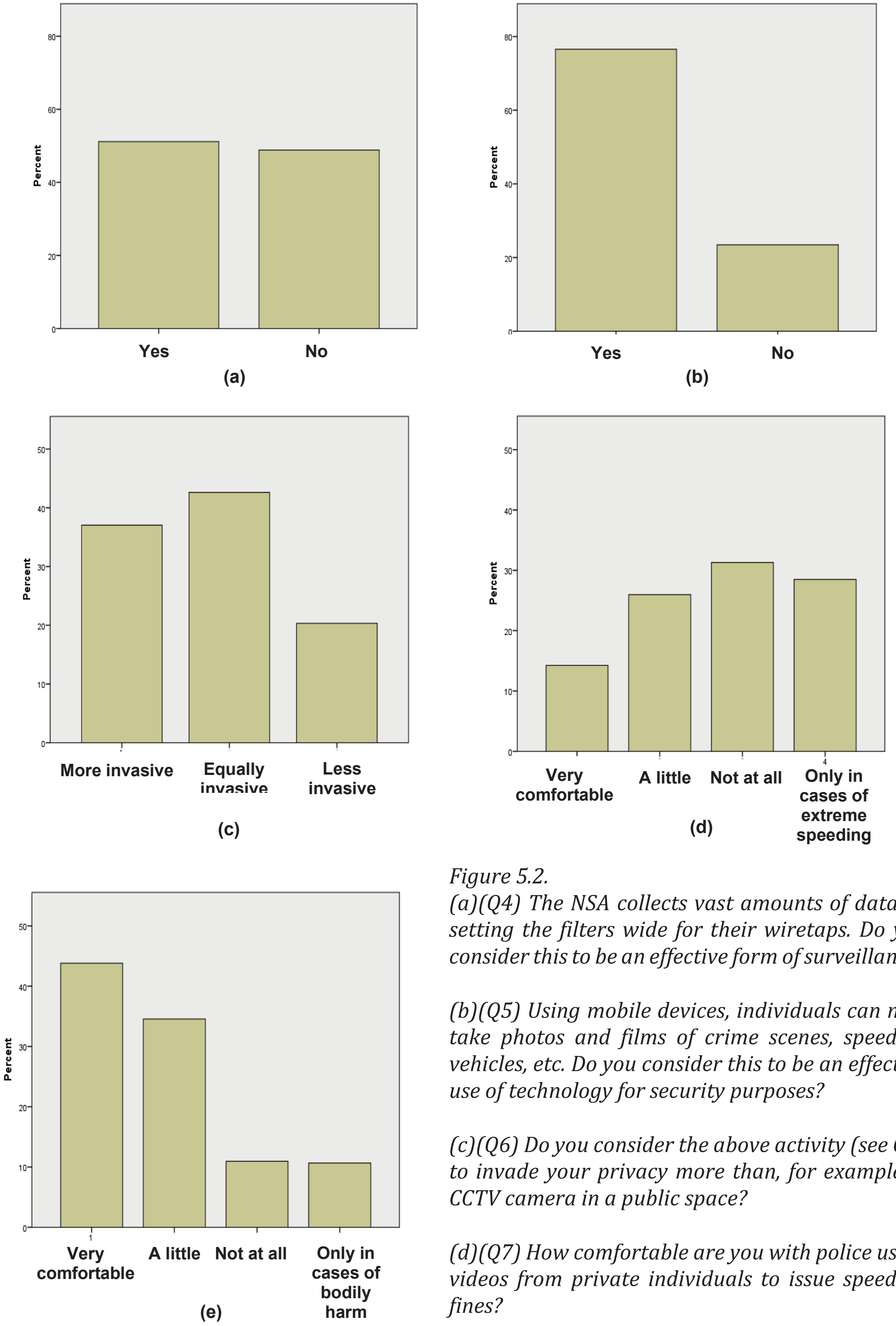


Figure 5.2.
(a)(Q4) The NSA collects vast amounts of data by setting the filters wide for their wiretaps. Do you consider this to be an effective form of surveillance?

(b)(Q5) Using mobile devices, individuals can now take photos and films of crime scenes, speeding vehicles, etc. Do you consider this to be an effective use of technology for security purposes?

(c)(Q6) Do you consider the above activity (see Q5) to invade your privacy more than, for example, a CCTV camera in a public space?

(d)(Q7) How comfortable are you with police using videos from private individuals to issue speeding fines?

(e)(Q8) How comfortable are you with police using videos from private individuals to investigate crime?

5.4.3 Cost

Perceptions of effectiveness and expectations of security influence how much money citizens are willing for the government to spend on surveillance technology; attitudes related to privacy may also influence attitudes toward cost.

> *H8.* *Those who find NSA wiretaps to be effective place less importance on the cost of surveillance technology, and those who find NSA wiretaps to be ineffective place more importance on cost.*
>
> *H9.* *Those that expect the government to prevent every possible terrorist attack place less importance on the cost of surveillance technology.*
>
> *H10.* *Those that place more importance on the cost of surveillance technology are more likely to be bothered by potential AIVD and NSA access to their online activity.*

Responding to how important the cost of surveillance technology is in relation to security [Figure 5.3, graph (a)], 31% said that cost is “very important” (that is, surveillance technology should only be deployed if the costs are reasonable), while 56% said it is “somewhat important” (surveillance technology that improves security can be somewhat costly). Only 13% responded that cost was “less important,” and that any surveillance technology that improves security should be deployed (*M* = 1.82, *SD* = .643). Here again, a difference in variance between students and parents manifested itself, with parents having a wider distribution (*p* = .038): 46 students (*M* = 1.87, *SD* = .582) and 45 parents (*M* = 1.78, *SD* = .704).

Respondents were pragmatic in not expecting the government to be able to prevent every possible terrorist attack [only 10% said yes – see Figure 5.3, graph (b)], and in their reaction when a terrorist attack occurs – only 8% said their response is that the government should have been doing a better job and prevented it [Figure 5.3, graph (c)]. A majority (56%) said that their reaction when a terrorist attack occurs is that the government is doing the best they can, while the remainder (37%) said their reaction is that something should be done to prevent this from happening again.

To test hypothesis *H8* we calculated the correlation in the responses to the question concerning the effectiveness of NSA surveillance technology and that investigating the importance of the cost of surveillance technology. That is, are those who considered NSA surveillance to be effective also willing for more money to be spent on surveillance technology. Contrary to our hypothesis, no significant correlation was found. However, the expected correlation between privacy and cost did appear (hypothesis *H10*), with those who were more bothered with AIVD surveillance of their online activity being more likely to say that cost is very important and surveillance technology should only be deployed if costs are reasonable (*r(91)* = -.280, *p* = .007). Interestingly, though, there was only a significant correlation between the cost question and the question related to domestic intelligence, not between cost and foreign intelligence.

A final correlation related to cost was identified: those who expected the government to prevent every possible attack were also more likely to consider cost as very important and that surveillance technology should only be deployed if the costs are reasonable (*r(91)* = .222, *p* = .035). This is contrary to hypothesis *H9*, which anticipated that those who expect the government to prevent every possible attack would be more likely to consider cost to be less important and that any surveillance technology that improves security should be deployed.

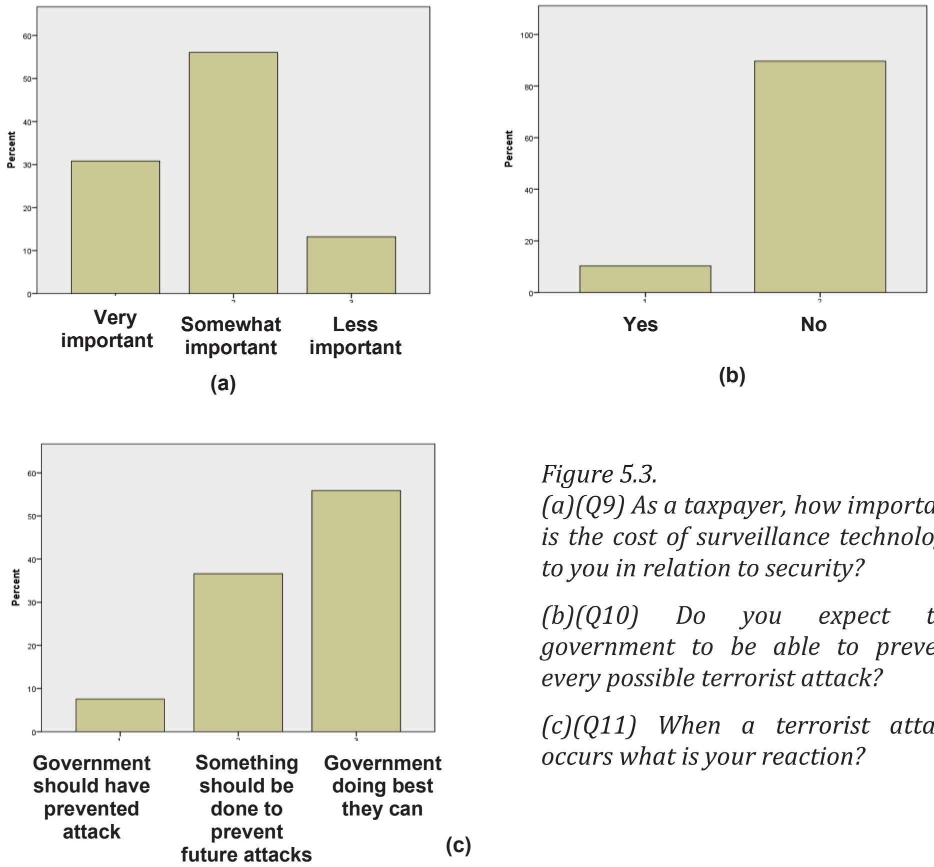


Figure 5.3.
(a)(Q9) As a taxpayer, how important is the cost of surveillance technology to you in relation to security?

(b)(Q10) Do you expect the government to be able to prevent every possible terrorist attack?

(c)(Q11) When a terrorist attack occurs what is your reaction?

5.4.4 Minor vs. non-minor students

In 2014, there were two groups of students surveyed – one group were students in the SSJ minor and the other group were not in the minor. It was anticipated that there would be some differences in the responses of these two groups. Those in the minor studied security and privacy issues surrounding surveillance technology, which we anticipated would make a difference in their responses.

> *H11. There will be a difference in the responses of SSJ minor students vs. students not in the minor.*

This turned out to be the case with several of the questions. Among the 68 SSJ students far more found private mobile devices recording crime and speeding to be effective for security purposes than not (82% found it to be effective, 18% not). A smaller majority of the 43 non-SSJ students found it to be effective (65%). According to a z-test for two population proportions (two tailed), the difference between 82% and 65% is significant (p=.042). The reverse was found regarding the effectiveness of NSA surveillance technology with 60.5% of non-SSJ students considering it to be effective, while only 46% of SSJ students considered it effective, although this difference is not significant (p=.136). A significant difference between the two groups was found regarding whether or not they expected the government to be able to prevent every possible terrorist attack (p<.001). A quarter (26%) of the students not in the minor expected the government to stop every attack versus only two students (3%) in the minor. The difference in means to responses to the next – and related – question mirrored these responses. Nine percent of SSJ students said the government should have been doing a better job versus 16% of non-SSJ students, although this difference is not significant (p=.262). The opposite end of the spectrum for this question was that the government is doing the best it can – an occasional terrorist attack will always get through. Sixty-two percent of SSJ students gave this answer, while only 40% of non-minor students gave this response, which is significantly different (p=.0394). Interestingly, differences that were expected in views of privacy invasion related to ISPs, the AIVD, and the NSA, did not appear.

5.4.5 Qualitative data

Respondents were asked to give free text responses as to why they considered NSA wiretapping surveillance technology with wide filters to be effective or not. We analyzed these responses by categorizing them according to themes. Strikingly, one of the most commonly stated reasons – a lot of data are collected – was given as an explanation both for why the technology was effective and ineffective. This was the number one reason that respondents gave for why they considered NSA wiretapping technology with wide filters to be effective – this technique yields a lot of information (30% gave this response). These answers stated that it is necessary to monitor all this information in order to find the crucial information you are seeking – only by monitoring the haystack will you find the needle. For this same reason many considered the technology to be ineffective – too much data are collected to be effective (15%). Two particular responses that underline the same reason given for both sides of the argument were:

> *Only by collecting vast amounts of potentially relevant data, will you end up with having that one phone call on tape that you happen to need.*
>
> *It is like looking for a needle in a haystack.*

The second response implies that the haystack should not be collected as it is so much information just to try to identify one small needle. While the first response argues that the only way to find that needle – that one important piece of information – is to collect the haystack.

Additional frequently given responses for why this technology is effective were that collecting information with wide filters allows the NSA to find what or who it is looking for, and that it is a preventative measure against crime and attacks. Interestingly, preventing crime and attacks was the 2nd highest reason given for effectiveness in 2014 and 2015, but in 2016 not one person gave this as a reason for effectiveness.

Other reasons people gave for considering the technology to be effective was that it allows the NSA to see everything and it identifies and predicts patterns. Several gave security as a reason, with some saying that anything that might increase security should be employed: "if there's a 1% chance that it would stop a terrorist attack or something like that, it is worth it." A small number reported that it was fast and easy and one person said the technology was cheap.

A number of respondents gave qualified "yes" responses. That is, this technology is effective, but… It is effective, but not ethical, or it is effective, but it invades privacy, or whether it is effective or not depends on how it is used or on the analysis of the data. Others said it was maybe or a little effective.

The first reason for ineffectiveness was that it invades privacy (18%). The second reason was that it collects too much data. The third reason given for it not being effective was that it yields a lot of unnecessary data (13%). (Note that we created two categories –

"too much data" and "unnecessary data" – but if the two were combined "too much data" would also be the number one reason for ineffectiveness.)

Other reasons that a significant number of people gave were that targeted surveillance is more effective, that criminals can work around filters and wiretaps, that it takes too much resource (time, money, and manpower), and that it makes it more difficult to find the important data. A smaller number of people said that such a large amount of data cannot be effectively analyzed, that this technique is not productive, and that it yields irrelevant information which results in false leads.

5.5 Discussion

The results from the three questions treating how bothered people were by private companies, the AIVD, and the NSA may point to matters of trust and politics. While not having all the capabilities of the NSA, the AIVD appears to be fairly advanced in their surveillance techniques and in cooperating and sharing information (Van Riezen and Roex 2012, Derix et al. 2013, Kraan 2017). Given the AIVD's capabilities, it does not seem likely that the difference in response toward the two agencies is related to surveillance powers. Rather, the fact that fewer respondents were bothered by the AIVD than by the NSA may indicate more trust in the Dutch AIVD, which would be expected – one would trust one's own country's intelligence agency more than a foreign one, particularly with all the bad press for the NSA around the Snowden leaks. The difference in variance between the AIVD and the ISPs might point to the AIVD being more of a political issue, and thereby evoking stronger feelings in both directions. More people being "not at all" bothered by the AIVD (28%) than by ISPs and internet companies (19%) could indicate more trust in the AIVD, or could again point to a difference in trust in domestic vs. foreign as some internet companies are foreign and even based in the U.S. (e.g. Google, which was mentioned in the survey as an example of an internet company). An alternative explanation could be that people feel they have no choice but to accept ISP access to their data. Based on their decision to accept this reality, people will adapt their values to view their decision positively (Adams 2014). Thus, they would be only a "little bothered," as most respondents were. All the post-Snowden discussion around privacy may also give people the impression that, contrary to ISP access, they do not have to accept AIVD access to their data, but can limit it through the democratic process (in reality, this could also be done for ISP access).

The correlations we discovered between responses related to effectiveness and privacy – e.g. those who considered NSA technology to be effective were also less bothered by potential NSA access to online activity, and vice versa – are similar to a finding of Pavone and Degli Esposti (2010) that people tend to be divided into two groups of being either trusting or concerned. Their study found that those who were trusting (of the institutions concerned and the technology's legitimacy) believed both that the surveillance technology was effective in increasing security and that their privacy was not infringed upon. Those who were concerned saw the technologies as an invasion of privacy without increasing security. Our findings suggest that those who are concerned about privacy (and perhaps see no added security value) are concerned regardless of who is looking at the data. And conversely, those who do not see a privacy invasion hold this view regardless of the institution involved. The Pew surveys cited earlier in this paper indicate American youth (ages 18-29) are more concerned about privacy and more disapproving of NSA surveillance programs than older adults. These reports, however, did not look for correlations between responses related to privacy and effectiveness. Our findings may suggest that the categories of being trusting or concerned, and therefore viewing surveillance technology as either both effective and non-privacy invasive, or the reverse, may be age-related; that students are not yet divided into groups of being trusting or concerned across the board, while their parents are. This could be because students are still forming and developing their views on issues and in politics, and are, as yet, not polarized. Their parents, on the other hand, have well-formed views and furthermore, some of these questions around security may be politicized, resulting in more polarization. This could also be an explanation for the difference noted above in the spread of responses between students and parents for Q1, Q2, Q7, and Q9.

As multiple authors have found (Pavone and Degli Esposti 2010, Van den Broek et al. 2017, Anderson 2015), and as these results suggest, trust is a core issue in surveillance. Trust in government institutions appears to equate to both trusting them with one's privacy and trusting that their surveillance technology is effective; distrust results in the reverse. The differing results according to age perhaps also speak to trust in political parties – a trust or distrust in the current governing political party.

At first review there was no correlation between effectiveness and cost. However, this might have been due to the effectiveness question zoning in on NSA surveillance technology, and the respondents being Dutch, and there, therefore, being no link between their taxpayer money and NSA technology. This interpretation is supported by the correlation found between cost and privacy as concerns the AIVD, while no correlation was found between cost and privacy related to the NSA. Further, there was a correlation between cost and expectations that the government would prevent every terrorist attack.

Expecting the government to prevent every possible terrorist attack points to respondents' expectations of effectiveness. The correlation this study discovered – that those who held this expectation were also more likely to consider the cost of surveillance technology as very important – is noteworthy. Although the sample size in this case is too small to make generalizations (the cost question only appeared in 2016), it merits further exploration to examine whether the public holds similar views related to cost and effectiveness as related to privacy and effectiveness/security (Van den Broek et al. 2017). That is, citizens do not treat the two as a trade-off, but as two independent elements, which both need to be attained.

Expectations that the government prevent every terrorist attack were significantly higher among non-SSJ students vs. SSJ students. One possible explanation is that, as a result of their studies, students of the minor were more conscious of the difficulties of preventing every possible terrorist attack, leading them to more realistic expectations. Of course, it is also possible that students who already held these views were more inclined to take the minor, and thus the differences are not the result of the studies. If this difference is a result of their studies, however, it is noteworthy that these studies did not significantly influence results in views on privacy and the effectiveness of NSA surveillance technology.

Among the whole respondent population, 56% said that their reaction when a terrorist attack occurs is that the government is doing the best they can. This majority could reflect a cultural aspect (e.g. the Dutch are a pragmatic people) and/or could be a result of there being no recent terrorist attacks in the Netherlands.

The contrast of respondents being considerably less comfortable with private videos being used to issue speeding fines than to investigate crime could indicate a consideration of more serious offenses meriting more invasive measures. An alternate explanation is that people are more averse to surveillance that touches them personally versus surveillance that does not concern them – i.e. we are probably all guilty of speeding, while few have committed a crime.

Limitations of this study include that its participants were exclusively in the Netherlands. Some of these findings, therefore, might only hold true for the Dutch. One study examining Dutch attitudes toward privacy and surveillance immediately following the Snowden leaks found them to be similar to attitudes in the U.K. (Mols and Janssen 2017). Obviously, many national factors can influence attitudes toward privacy and surveillance, such as surveillance law, history, scale of terrorist attacks, etc. Future research could compare this paper's findings with similar studies in other countries.

Survey questions on the effectiveness of surveillance technology focused on one type of technology, and on private mobile device use in two scenarios. Other types of surveillance technology or different scenarios might yield different results. Although the results related to SSJ minor students vs. non-minor students indicate significant differences in some areas and not others, further studies are needed to confirm these findings, as well as to determine to what degree student findings across all years might be influenced by most students being in security studies. As previously mentioned, the cost question only appeared in one year. The results related to participants' views on the acceptable cost of surveillance technology merit further research to investigate if the lack of trade-off between cost and effectiveness holds true with a larger sample size.

5.6 Conclusion

Surveying the public on the effectiveness, cost, and privacy implications of surveillance technology yielded some noteworthy results. This study's findings support other studies which have found that people tend to be either trusting (believe surveillance technology to be effective and non-invasive) or concerned (find surveillance technology to be ineffective and invasive), and do not engage in a privacy-security trade-off. This paper found that this non-trade-off extends from privacy-security to privacy-effectiveness. It also suggests that the public being divided into groups of either trusting or concerned is age/generation dependent – students are not yet either trusting or concerned across the board, and show no correlation between effectiveness and privacy. Another key finding is that as with privacy-effectiveness, also with effectiveness-cost – the public does not engage in a trade-off here either, but rather expects both to be achieved.

Only one significant difference revealed itself between SSJ minor and non-minor students, with the first group having lower expectations of effectiveness. Lastly, respondents gave the same top reasons for both why NSA surveillance technology was effective and ineffective – collection of a lot of data.

This study's findings contribute to the surveillance debate in investigating these questions and correlations for the first time. They support recent literature in the privacy-security debate, which shows that the public does not engage in trade-offs between the different values involved, but rather, wants it all. The results merit further investigation, particularly as regards the age-correlation. It is a significant finding that groups of trusting or concerned people appear to be age-related. The differences between parents and students, and between students in and outside the minor also suggest that this is an age when views on security matters are shaped and with time are increasingly solidified. To have real and inclusive dialogue, policy-makers should move away from a trade-off mentality and engage in dialogue with the public sooner rather than later.

Chapter 6
Fettered by Bureaucracy[1]

6.1 Introduction

Surveillance is a necessary function of any intelligence agency. Through surveillance intelligence services strive to know the intentions and capabilities of potential adversaries without themselves being known. Surveillance is also, however, a controversial topic with some stakeholders railing that certain surveillance measures are too invasive, while others declare they are necessary for effective security. Intelligence agencies seek to carry out their mission of defending the nation, while parliamentary bodies pass laws to ensure intelligence services are accountable and kept in check. The public demands its privacy be protected, at the same time expecting effective surveillance practices to deliver security. It appears, however, that no one is assessing if, in fact, the surveillance technology employed is effective in achieving its security goal.

Previous studies reveal that effectiveness is minimally treated by both intelligence practitioners and oversight bodies (Cayford and Pieters 2018, Cayford et al. 2018). Additionally, intelligence agencies' goals often clash and practitioners appear to pass the buck on proportionality, asserting that proportionality judgments are not theirs to make (Cayford and Pieters 2018). These and other findings beg for explanation. Why do these institutions operate in this way? Why is effectiveness evaluation of surveillance technologically so absent? Is it impotent politicians unable to develop policy, poor oversight, or government officials who just do not care if surveillance is effective? This article seeks to dive below the surface to seek out the underpinning explanations.

These institutions operate in the context of a democracy, ultimately being beholden to the people to carry out specific mandates, while following the law and limitations laid down by a democractic government. A possible explanation to the questions posed above is that democratic bureaucracy is itself to blame. A bureaucracy that is required by we, the people. The manner in which the effectiveness of surveillance technology is handled could be related to the bureaucratic constraints under which these institutions operate. That is, the bureaucratic system itself, necessary for the operation of democratic government, may be a root cause. To test this hypothesis, this article analyzes James Q. Wilson's book, *Bureaucracy* (1989), in the context of evaluating the effectiveness of

[1] This chapter has been accepted for publication with *Intelligence and National Security*.

surveillance technology. This study looks at how government practice in surveillance can be interpreted in light of Wilson's observations and explanations on the workings of bureaucratic government.

This article proceeds with a brief review of key findings on the evaluation of the effectiveness of surveillance technology, followed by a presentation of pertinent elements from Wilson's observations on bureaucracy. These elements are then analyzed in the context of the key effectiveness findings. Finally, a discussion section considers differences between American and European bureaucracy and draws on additional theories and literature from economics and governance to further develop the article's findings.

6.2 Previous research

The lack of effectiveness evaluation is reflected in the broader surveillance discussion with existing surveillance literature minimally touching this subject. Authors have written about the intersection of privacy and modern technology and the challenges that arise (Greenwald 2014, Berghel 2013, Morgan 2014, Monahan 2016, Bigo et al. 2013), on mission creep (Regan and Monahan 2014, Monahan and Palmer 2009), and the success of intelligence oversight (Ford 2006, Ott 2003, Zegart and Quinn 2010, Dietrich 2016). Existing studies on effectiveness tend to be counterterrorism centered (Lum et al. 2007, Van Dongen 2009 & 2015, Van Um and Pisoiu 2011, Jonas and Harper 2006) or focused on specific technology (Lingel et al. 2012, Willis et al. 2010, Tsvetovat and Carley 2006). Effectiveness has also been discussed in the context of cost-benefit analysis (Stewart and Mueller 2011, Mueller and Stewart 2011).

This article takes its place in this broader discussion by examining the overall picture of the government's handling of effectiveness in relation to surveillance technology and its root causes.

The authors' larger research study, of which this article forms a part, focused on intelligence agencies in the U.S. and U.K., and the surveillance technology they employ (Cayford and Pieters 2018, Cayford et al. 2018, Cayford et al. 2019). It examined the question of how the effectiveness of surveillance technology is treated by stakeholders, as well as how considerations of cost and proportionality figure in their assessment. The term "surveillance technology" included a broad range of technologies, including drones, satellites, wiretaps, cameras, etc. In the data analyzed the majority of the technology under discussion was that dealing with communications data, such as surveillance systems monitoring internet activity, phone calls, etc.

Effective in this study was defined as, "an impact that is desirable and can be observed as contributing toward the sought-after security goal." This differs from *performance* which refers to the technology's ability to function correctly. Effectiveness is never determined in a vacuum. Officials determine to use or not a particular technology, not based solely on strict effectiveness, but on considerations of cost and/or proportionality. Because these factors play a role in any evaluation, how they are treated was also analyzed. This study refers to these three elements as overall effectiveness (composed of strict effectiveness, cost, and proportionality).

According to one understanding of proportionality, effectiveness could be considered as included in assessments of proportionality. That is, proportionality is an assessment of balance between the expected harms (privacy intrusion, economic cost) and benefits (security and economic gains). This is particularly pertinent as regards the U.K., where law enforcement is required to perform a proportionality assessment prior to the deployment of any surveillance technology. Using this understanding it could be argued that effectivess is already assessed, in the context of security benefits.

This article and its greater study does not take this approach for several reasons. Firstly, effectiveness is an important element and deserves to be evaluated on its own. The danger of considering it only within the context of proportionality is that it is immediately obscured by the question of privacy. It is no longer a question of whether it is effective, but whether it is effective vis-à-vis privacy. This risks drowning out many important questions. Assessing effectiveness is complex, raising questions such as how is it evaluated, against which measures is it assessed, who is judging and determining effectiveness, what is the threshold of effectiveness, are different measures used for strategic versus tactical operations, and how to evaluate technology used for multiple purposes. These questions risk being lost or over simplified if effectiveness is only treated as a security benefit in the context of proportionality. Secondly, effectiveness and proportionality are potentially conflicting elements. They are both desirable and sought after goals, and yet a technology may be effective, but not proportional, and vice versa. As such, effectiveness should be evaluated independently before being assessed in the context of proportionality. Finally, focusing strictly on effectiveness necessitates a security goal. To determine effectiveness a technology must be measured against the specific goal it is meant to achieve. Addressing effectiveness only by including it in assessments of proportionality makes it easier for the security goal to be overlooked and to gravitate toward broad and general statements, such as a program "increases security" or this is an "effective agency."

The larger research study focused on three intelligence agencies in the U.S. and the U.K. – the NSA, CIA, and GCHQ. These countries and agencies were chosen due to the attention they received following the Snowden leaks and the consequent availability of data. The study investigated how three different groups of stakeholders view and evaluate the effectiveness of surveillance technology. Firstly, statements of intelligence practitioners were examined to gain insight into their criteria for determining effectiveness. The second group studied was oversight bodies and their determinations of effectiveness. Finally, surveys were conducted to discover how the public views questions of effectiveness.[2]

The findings of these three papers evidences two differing perspectives regarding the effectiveness of surveillance technology: one recognizes a trilemma and that all three elements cannot be delivered at once; the other rejects a notion of trade-off. This difference of perspectives between the public and government agencies causes one to search for the source of this log jam. Further, one would expect government agencies to evaluate effectiveness and yet there is a remarkable lack of such assessment.

This article seeks elucidation for the following key findings:

(1) Effectiveness of surveillance technology is minimally addressed in intelligence agencies.
(2) Oversight bodies rarely evaluate effectiveness, but rather call on the intelligence agencies to do so.
(3) Intelligence agencies appear to have conflicting goals of performing effective surveillance, at low cost, with minimal privacy intrusion.
(4) Intelligence officials consider proportionality to be addressed by the law and not a decision that they themselves must make.
(5) Oversight bodies and intelligence agencies do not simultaneously address effectiveness, cost, and privacy, inherently recognizing a trilemma; the public, however, does not perform a trade-off but expects all three to be simultaneously delivered.

This article hypothesizes that these findings can be explained by bureaucracy and the restraints it places on government institutions.

6.3 Observing bureaucracy

To test this hypothesis, this article turns to James Q. Wilson's seminal work on the bureaucratic nature of democratic government – *Bureaucracy* (1989). In this book

[2] A more detailed summary of these three studies is included here in the version of this chapter submitted for publication. It is omitted here since such a summary is unnecessary.

Wilson examines how bureaucracy works, explaining why government agencies function the way they do, often in a seemingly inefficient and/or inept manner. This book was chosen due to Wilson being the most authoritative figure on bureaucracy. According to Google Scholar *Bureaucracy* has been cited 6478 times. Pietro S. Nivola of the Brookings Institute states that this book is "Still widely regarded as more or less the last word on the subject" (Nivola 2012). Additionally, what Wilson does mirrors what this paper seeks to do – explain why a government institution behaves the way it does. Wilson, of course, studies and explains the behavior of many bureaucratic offices, while this article focuses on five aspects in one area of bureaucracy. The advantage of this is it allows for the examination of intelligence agencies against the key features of bureaucracy, as opposed to perhaps concluding that any findings are bureaucratic features unique to intelligence bodies. One potential limitation to this book choice is that Wilson focuses on U.S. bureaucracy, while our research examines both U.S. and U.K. government agencies. However, the five key findings cited above apply to both the U.S. and the U.K. Thus, if this current article finds that Wilson's work helps explain these previous findings vis-à-vis the U.S., this could hold true for the U.K. as well. This point is further discussed in the Discussion section of this article.

Reading through Wilson's enlightening book, naturally, not every point, although interesting, is relevant to this present study. Specific elements were chosen from *Bureaucracy* for analysis here, based on their ability to shed light on the evaluation of surveillance technology in government agencies. This section presents these factors, and the following section analyzes them in light of effectiveness evaluation of surveillance technology.

Goals

Wilson observes that government agencies are likely to have "general, vague, or inconsistent goals" (Wilson 1989, p.26). For example, the Department of State's goal is to "Promote the long-range security and well-being of the United States" (p.32). And within the Department of Housing is the goal to "Develop viable urban communities by providing decent housing and a suitable living environment" (p.32). These general, vague, or inconsistent goals exist because people disagree not only on the meanings of words (what does "security" or "advantage" mean?), but also on how these goals should be attained. Should other goals be sacrificed to reach these ones? For example, should "decent" housing be provided regardless of the cost (p.33)?

Wilson argues that in the absence of clear and specific goals, employees' work will be shaped by circumstances (as well as by personal beliefs and experiences). Circumstances include the clients' behavior and the tactics available to the worker. Regardless of what the agency's goals are, these factors will shape what the workers will do. Even with clearer goals, the situation can define a worker's tasks if a certain way of performing the job seems easier (Wilson, p.42). Wilson cites the Occupational Safety and Health Administration (OSHA) as an example. OSHA was created to protect worker safety and health. At the time of its creation, dangers to worker health (e.g. scores of workers dying due to exposure to hazardous chemicals) was considered to be a greater threat than occasional injury due to safety issues. But in the years since its formation OSHA has focused much more on rules surrounding safety rather than health. This is because it is easier. It is easier to assess the cause, cost, and solution to a worker falling and breaking a leg, than to do the same for a worker who develops cancer after years of working at a chemical plant (p.42).

Government agencies are brought into existence to serve certain goals. However, an agency is expected not only to serve its primary goal but also *contextual* goals. These define the context, or state of affairs, the agency must maintain while seeking its primary goal. For example, a police department's primary goal is to prevent crime and arrest criminals, but it must also protect the rights of the arrested person, ensure its records remain confidential, and provide health services to those arrested (Wilson, p.129).

In addition to an agency's primary and contextual goals, the government overseeing and directing it often has a plethora of its own objectives. Discussing the case of the U.S. Postal Service, which was previously under Congress and the president, Wilson points out that Congress' goal was not simply to speedily deliver the mail at the lowest possible cost. Rather, it wanted to satisfy different categories of mail users, as well as constituency demands to maintain many small post offices, address wage increase demands, and respond to public dissatisfaction with the mail service.

> *Congress could not provide a consistent rank-ordering of these goals, which is to say that it could not decide on how much of one goal (e.g. keeping prices low) should be sacrificed to attain more of another goal (e.g. keeping rural post offices open). This inability to decide is… the inevitable consequence of Congress being a representative body whose individual members respond differently to different constituencies (Wilson, p.125).*

Constraints over tasks

Government agencies operate under a considerable number of constraints: their revenues cannot legally benefit the organization's members, they must serve goals that the organizations themselves have not chosen, and they cannot allocate resources according to their preferences (Wilson, p.115).

All these constraints and the contextual goals mentioned above have an effect on the *management* of public agencies. Managers have reason to be more concerned with constraints than tasks. This means that the process becomes more the focus than the outcome. While outcomes are often "uncertain, delayed, and controversial[,] procedures are known, immediate, and defined by law or rule. It is hard to hold managers accountable for attaining a goal, easy to hold them accountable for conforming to the rules" (Wilson, p.131).

Constraints also result in fairness becoming more important than efficiency. This is because fairness is easier to judge than efficiency. For example, it is easier to judge whether every student got the same textbook than it is to judge whether the students were educated (Wilson, p.132).[3]

Standard operating procedures (SOPs) are a very tangible result of constraints. In order to ensure that constraints and contextual goals are not violated, SOPs are put in place. According to Wilson, these rules are a way to hold agencies accountable to constituencies, punishing those who upset the constituencies (Wilson, p.133).

Visible output and invisible outcome

Wilson classifies government agencies into four groups according to whether the employees' activities (output) and the results of those activities (outcome) is observable or not to managers. We argue that two of these categories apply to intelligence agencies – coping organizations, in which neither the output nor the outcome is observable to managers, and procedural organizations, in which outputs are observable, but not outcomes. With regard to agents in the field an intelligence agency is a coping organization, with neither the agent's daily actions nor the outcomes of his/her actions visible to management. Intelligence analysts, however, work in an office under the observation of the manager. But it is still difficult for the manager to know whether the analyst's actions gave the country strategic advantage over its adversaries. According to Wilson, in procedural organizations, what becomes more important is *how* the

[3] Wilson speaks of *efficiency*. Although *efficiency* and *effectivenes*s have two different meanings, in this context effectiveness could be used in place of efficiency. That is, it is easier to judge matters of fairness (every student getting the same textbook) than to judge whether the education system is effective in educating students.

employees do their jobs, rather than whether the execution of the job results in the desired outcomes. When the results of the work are unseen or hard to determine then it is difficult to convince others that what the agency is doing really works or is effective. In this case, rules take on more importance – SOPs become "pervasive" – while managers seek to convince political superiors that their agencies are faithfully following the rules (Wilson, p.164).

Rules without trade-offs

According to Wilson, delivering a public service can either be improved by rules or contracts. It is rules and not contracts that pervade agencies because government institutions find rules to be more rewarding. A rule appears to immediately respond to a constituent's grievance. And it usually does not have to be reconciled with other rules, freeing political institutions from having to make difficult choices among contending goals. Government bodies imposing the rules do not need to address issues of feasibility. A contract, by contrast, has to reconcile at least some major tradeoffs – time and money, cost and quantity. Political actors, however, do not consider trade-offs. They want to see all of their interests protected (Wilson, p.363). And the government itself "is required to act as if all preferences can be accommodated simultaneously" (p.364).

Political environment

Wilson points out that much of the difficulty of bureaucracy arises from the fact that the government is a democracy. It is not simply a management problem, but a governance problem (Wilson, p.376). A government agency must maintain political support. A typical agency, however, is not in an easy position to gain political support as it often:

> *must do something that is unpopular (e.g., collecting taxes) or difficult (e.g., managing foreign affairs) and that a half dozen other agencies are doing (e.g., gathering intelligence or catching drug dealers), and it must do these things under the watchful and critical eyes of countless subcommittees, interest groups, and journalists. It faces inadequate budgets, complex tasks, several rivals, and many constraints (p.181).*

To "fix" these problems – grant a simple and clear mission, minimize constraints, judge officials on outputs rather than inputs, etc. – political actors would have to function contrary to their own interests. They would have to relinquish seeking to expand their own influence, turn down influential constituents, and seriously consider the feasibility and political popularity of any proposed new program. This scenario seems highly unlikely both because politicians have no incentive to bring it about and because "there are certain tasks a democratic government must undertake even if they cannot be performed efficiently" (p.376).

The bureaucracy of government is a result of the political system and of we, the people, who require it to be so (Wilson, p.133). Democratic government must gather intelligence and perform surveillance, it must respond to the demands of its constituents and their differing goals and preferences, and it is required to operate under numerous constraints imposed by the people. Constraints rather than the tasks themselves become the focus, fairness takes priority over effectiveness, and circumstances rather than goals shape employee's work. Bureaucracy comes part and parcel with democracy and has a direct effect on effectiveness evaluation.

6.4 Analysis

These identified elements from *Bureaucracy* shed light on some of the key findings of this larger study's research on the evaluation of the effectiveness of surveillance technology. This section interprets these findings in light of Wilson's observations on bureaucracy. A summary of previous research findings explained by Wilson's concepts is found in Table 6.1.

Goals (in intelligence agencies)

Applying Wilson's concepts to key finding #3 (Intelligence agencies appear to have conflicting goals of performing effective surveillance, at low cost, with minimal privacy intrusion) one observes that intelligence agencies are also clearly affected by the goals of elected officials in the legislative body. Because politicians have a vast constituency and seek to please all their constituents, they do not just have one goal, but many conflicting goals – effective surveillance technology, low cost surveillance, protected privacy. As a collective body, elected officials cannot decide which goal has priority, and which goal should be sacrificed at the cost of another. The result is that intelligence agencies themselves pursue conflicting aims.

In recent years, intelligence agencies' surveillance practices have been under fire on grounds of privacy and proportionality (e.g. Greenwald 2014, Bergen et al. 2014, Berghel 2013). Intelligence officials consider proportionality to be addressed by the law and not by themselves (key finding #4). Wilson's observations bring understanding to this finding – protecting innocent civilians' privacy is not a primary goal of intelligence agencies – it is a contextual goal. They cannot perform surveillance willy-nilly – the condition under which they must keep the nation secure is that their surveillance must be carried out in a proportionate manner. This sheds some light on how intelligence practitioners treat proportionality. It is not their primary goal because protecting the privacy of citizens is not the agency's mission. But it is a contextual goal under which they must carry out their primary goal of securing the nation. Intelligence practitioners view questions of proportionality and privacy as addressed by the law and not by themselves. That is, the law and oversight bodies establish parameters for what is proportional, and they act within these limits. For intelligence practitioners, their job is to provide intelligence, while operating within the legal framework, and not to make judgments on proportionality. This is in keeping with proportionality being a contextual goal. A further point to consider is that if this was an additional primary goal it would conflict with the goal of securing the nation, which would arguably lead to agency paralysis.

Observations on goals help explain why intelligence officials' handle privacy and proportionality as being treated by the law and not by themselves – it is a contextual, not primary goal. They also reveal the effect on intelligence agencies of elected officials having conflicting goals – no one goal is given priority. Because proportionality is a contextual goal and not a primary one, intelligence services avoid paralysis. But the fact that the contextual goal still conflicts with the primary goal nonetheless introduces an aspect of impossibility of attempting to satisfy conflicting goals.

Effectiveness minimally treated

Intelligence officials were found to minimally treat the question of effectiveness (key finding #1). Intelligence agencies appear to not formally evaluate the effectiveness of surveillance technology; evaluations that do consider effectiveness focus on intelligence agencies as a whole, rather than particular surveillance programs or technologies. Compliance reports, on the other hand, do clearly exist. Intelligence agencies must regularly report if there have been any breaches of compliance, what these breaches are, and what they have done to remedy the problem (Cayford and Wolters 2018).

Several of Wilson's observations explain this phenomenon. The first explanation concerns circumstances shaping employees' work, particularly when goals are vague or general. In this scenario workers will perform a task in the way that seems easiest. This principle can be applied to intelligence agencies in a broad sense. Effectiveness is difficult to evaluate. Is the surveillance technology considered effective if it provides X amount of information? How is 'X' determined? Or if it only leads to one key piece of information? Did that information play a key role in dismantling a criminal organization? Was the information crucial to informing policy? It is easier to judge legal issues and cost than if, or to what degree, particular technology has contributed to effectiveness. Whether a surveillance program has stayed within budget or is too costly, or whether its use has complied with legal requirements is much easier to judge than whether or not it has contributed to meeting a said security goal.

Similarly, the constraints under which managers in government agencies operate result in fairness becoming more important than effectiveness because it is easier to judge than effectiveness. This aspect of managers tending toward the more easily measured element applies to intelligence agencies. In managing surveillance programs, it is easier to pay attention to following rules (law) and budget than to judge effectiveness. For both workers and managers judging effectiveness is more difficult than paying attention to rules and budget. Consequently, effectiveness is minimally treated.

A third explanation relates to Wilson's classification of four types of government organizations. Intelligence agencies move between two of the categories – coping and procedural – depending on the type of employee. Analysts' work is observable to managers, while agents' work in the field is not. This paper focuses on the work of analysts because the principle kind of surveillance technology discussed in the greater study was that dealing with communications data. Analysts are the employees that handle this data. Their administrators can observe how the analysts use the surveillance technology. Determining the outcome of their work or the effectiveness of using a particular technology, however, is more elusive. In some cases it may be clear that certain actions contributed to the successful outcome of a tactical operation. More often, however, it may be difficult to determine if the use of the technology resulted in the dismantling of a criminal organization. And for strategic operations it may be impossible to measure if the action resulted in strategic advantage or a change in government policy. What or how much of an outcome to attribute to a particular technology is difficult to evaluate. Wilson classifies agencies in which the employees' work is observable, but the outcome of their work is not, as procedural organizations. Because this greater study focused on surveillance technology used by analysts, for the purposes of this paper – that is, in this particular context – intelligence agencies are classified as procedural organizations.

According to Wilson, standard operating procedures (SOPs) pervade procedural organizations. This may be less the case with intelligence agencies since they are not exclusively procedural organizations, but this does explain why compliance reports seem to exist in abundance, while effectiveness reports do not. The use of surveillance technology, particularly in regards to internet-related surveillance, is observable and therefore is subject to procedural rules. Analysts must follow strict guidelines in their use of this technology. Compliance reports are produced on a fixed-term basis, documenting instances in which the rules were breached and correctives taken.

Minimal treatment of effectiveness in intelligence agencies is explained by the fact that both employees and managers tend toward the more easily measured elements of cost and legality, as well as by intelligence agencies being (in part) procedural organizations, in which analysts' actions are observable, but not the result of their actions, leading to an abundance of compliance reports.

Complex oversight

This study's research on oversight bodies found that intelligence oversight bodies rarely, if ever, evaluate effectiveness (key finding #2). Instead, they push the intelligence agencies to do so. Intelligence oversight concentrates on matters of cost and legality, calling out agencies and programs that have gone over budget or who have had breaches of law. But it does not evaluate effectiveness within these procedural organizations. Wilson points out that procedural organizations present an oversight problem. Oversight bodies can and do put rules and procedures in place to direct how tasks are carried out. But they do not help in evaluating how well the job has been done. For example, the U.S. Congress required the Marine Corps to change its training methods after determining the training was too abusive. But it could not evaluate which training method produced the best marines (Wilson, pp.245-246). This matches the findings of this greater study's research on intelligence oversight bodies: they were found to rarely evaluate effectiveness – or how well the job was done – themselves (Cayford et al. 2018).

This suggests that the difficulty with intelligence oversight is not only, as so many claim, that intelligence agencies operate in secret, but also that they are (in part) procedural organizations in which the outcome cannot be judged. It is difficult to oversee and fully hold accountable, an organization which performs tasks, the success of which is difficult to determine.

Operating under politics in a democracy

Discussions surrounding surveillance are well-known for the discourse of trade-off: if more effective surveillance is desired, some privacy must be given up and vice versa. This study's article on the public investigated its view of the concept of trade-off off between effectiveness and privacy, and effectiveness and cost. The findings suggest that people do not perform this kind of trade-off, but rather want effectiveness, cost, and privacy delivered simultaneously. All three elements appear to be given equal priority by the public (key finding #5).

It is, perhaps, sometimes forgotten that intelligence agencies are also influenced by politics. As Wilson points out, parliamentary bodies have many constituents with many differing views, and each politician wishes to satisfy his/her constituency. Politicians and officials seek to be reelected and to influence policy making. Consequently, Congress (or Parliament) itself has multiple goals, and it is unable to prioritize which of these goals is most important. In the case of surveillance technology, which goal should have priority – that the technology is effective, that it is low cost, or that it is proportional in its invasion of privacy? Congress cannot arrive at a unanimous decision. This may not be just a result of some people wanting effectiveness to be prioritized while others want privacy first, but of citizens wanting it all.

Further, as Wilson highlights, there are some tasks a democracy must perform, even if it cannot do so efficiently (or effectively). It must carry out intelligence operations using surveillance. In a democracy the people demand effective security, that the carrying out of that security not waste their tax money, that it is kept within certain bounds and not abused, and therefore that oversight is put in place, and that the public's privacy is not infringed upon at the expense of that security. Hence surveillance finds itself in an impossible trilemma of needing to simultaneously meet these conflicting demands as a result of being part of a democracy. A democratic government must perform surveillance and must strive to meet its constituents' conflicting demands. Congressional bodies are unable to prioritize one goal over another, both because they represent many constituencies and because the public appears to not prioritize one goal over another.

Research finding	**Applicable concept from Wilson**	**Interpretation of finding using Wilson's concept**
Intelligence officials consider proportionality as addressed by law	Contextual goals	*Proportionality is a contextual, not primary goal*
Conflicting goals in intelligence agencies	Conflicting goals of elected officials	*Legislative body cannot decide which goal has priority resulting in intelligence agencies having conflicting goals*
Effectiveness minimally treated	Workers choose easiest way to perform task Managers evaluate more easily measured elements In procedural organizations outcomes are difficult to evaluate	*Effectiveness is more difficult to evaluate than measuring accordance with rules and budget* *Outcomes of analysts' actions are unobservable resulting in compliance reports but not effectiveness reports*
Oversight bodies do not evaluate effectiveness	Procedural organizations are difficult to oversee – oversight says how job should be done, but does not evaluate if it is done well	*Intelligence oversight is unable to judge the success of agencies' tasks*
The public does not perform a trade-off between effectiveness, cost, and privacy	Congress has multiple, un-prioritized goals, due to representing many constituencies	*Surveillance finds itself in an impossible trilemma of needing to simultaneously meet three conflicting demands*

Table 6.1. Wilson's concepts applied to research findings

6.5 Discussion

Differences across the pond?

Wilson's work focuses on U.S. bureaucracy. This study has centered on both American and British intelligence agencies and their oversight bodies. This arguably leaves room for some of Wilson's observations to not be applicable to all of this research. The elements drawn from Wilson for this article, however, have more applicability on both sides of the pond than not. These countries are both democracies, meaning they are subject to the rule of the people and their institutions are influenced by politics.

In discussing accountability and transparency in a European context, Christiansen and Lodge (2016) employ Wilson's categorization of government agencies into four groups (production, procedural, craft, and coping). The authors examine three different types of government agencies - intelligence, flood defense, and food safety – across five European countries (U.K., Germany, Denmark, Sweden, Norway). They argue that with intelligence agencies it is difficult to measure outputs and outcomes. This is according to Wilson's categorization system. It is also one of the classifications we have given intelligence agencies – that is, a coping organization. In their analysis, the authors found that variation occurs related to the type of organization and its tasks, rather than according to the countries. This suggests that despite differences in government types, the agencies across these countries and their ways of operating are more similar than not. This finding, in addition to the authors' use of Wilson's categorization of agencies in a European context, supports the applicability of Wilson's work to the U.K. portion of our larger research study.

Wilson himself addresses differences between American and European bureaucracy. One difference he highlights is that European government agencies are beholden to fewer contextual goals than their American associates (Wilson, p.130). This research, however, analyzed the contextual goal of proportionality. Proportionality is a clearly stated contextual goal for British intelligence services (Cayford and Pieters 2018).

While the U.S. and the U.K. are both democracies, they are of different types. The U.S. is a presidential democracy while the U.K. is a parliamentary democracy. In a parliamentary democracy the authority to make and implement policy is concentrated in the hands of the executive. The prime minister and parliament are not rivals. In the U.S. the executive and legislative branches *are* rivals, each trying to limit the other's power while increasing their own, and serving as a check and balance to one another. In a presidential democracy Congress has power independent of the president, while

Parliament has little authority and cannot investigate agencies if the prime minister objects (Wilson, p.298).

British oversight bodies, then, are under the command of the prime minister, while the American ones are divided between the executive and legislative branches. One could imagine the Congressional oversight bodies to be more critical and strict than their British counterparts, as they seek to keep in check the executive branch, under whose direction the intelligence community lies. For example, one might suppose that Congress would conduct evaluations of effectiveness to “prove” or “disprove” that certain surveillance actions are effective. As this article underlines, however, this is not the case.

Conversely, one could imagine that the reason British oversight bodies do not evaluate effectiveness is that the intelligence agencies are beholden to the prime minister, who does not want to bring to light below par performance. According to this logic, however, all oversight bodies across the whole government would be pointless, as they would not be holding agencies accountable. Furthermore, the intelligence oversight arms of the Interception of Communications Commissioner and the Intelligence Services Commissioner (now grouped under one Investigatory Powers Commissioner) investigate and report on compliance and proportionality issues. These reports are annual, accessible to the public, and can be critical of transgressions committed by the intelligence services (Cayford et al. 2018). Therefore, it is not likely due to differences in government structure that effectiveness is not evaluated in the U.K., but more probably that Wilson’s observations about U.S. bureaucracy analyzed in this article are also applicable in the U.K.

Further dimension added by additional theories

Wilson’s observations reveal that the structure and nature of bureaucracy is a large factor in the lack of effectiveness evaluation of surveillance technology. Theories from other domains further explain why this lack exists. These theories provide explanation from other angles. Bureaucratic elements are perhaps the most fundamental reason; these additional theories provide increased depth as to why effectiveness evaluation is so elusive.

Intelligence agencies and oversight bodies do not seek to simultaneously evaluate and obtain the three conflicting goals of effectiveness, cost, and proportionality not only because of bureaucracy, but also because it is impossible. The trilemma concept from macroeconomics states that only two of three conflicting goals can be simultaneously achieved (Aizemann and Ito 2012, Obstfeld et al. 2004). Monetary independence, exchange rate stability, and financial integration are all desirable goals in open economies, yet it is impossible to simultaneously achieve all three. Therefore, policymakers must decide which one of the three they will give up. This same reality exists in the security realm, and is reflected particularly by oversight bodies. The impossibility of obtaining effectiveness, cost, and proportionality simultaneously leads to addressing a maximum of two goals at a time.

To avoid engaging in a trade-off, oversight bodies are compartmentalized to deal with specific issues. Steenhuisen (2009) found that oversight bodies distance themselves from trade-offs. This study's findings support this argument, concluding that intelligence oversight bodies are either created to deal with only one of the three elements of effectiveness or if they are tasked with overseeing all three elements, only evaluate one or possibly two at a time. Examples include the Privacy and Civil Liberties Oversight Board, which as the name suggests, focuses on protecting the privacy of Americans in matters of surveillance; the U.K. Intelligence Services Commissioner produces yearly reports on the legality of warrants issued; the Dutch oversight body, CTIVD (Intelligence and Security Services Review Committee), states explicitly that it does not evaluate effectiveness. This research found no oversight reports that dealt with evaluating the effectiveness, cost, and proportionality of surveillance programs simultaneously. Having a compartmentalized approach allows oversight bodies to avoid engaging in a trade-off and successfully address the issue at hand, rather than having to enter into the impossible trilemma. Oversight's distance from trade-offs is an additional feature of bureaucracy – one not specifically addressed by Wilson.

The concept of trade-off is also found in governance literature, which discusses values and value conflict. Graaf and Wal (2010) argue that effective governance and ethical governance clash; that public governance cannot reach its objectives while being "good." "Truthfulness, decency, and transparency do not characterize the spirit of effectiveness. What is more, infractions such as rule-bending, selective honesty, and the resetting of agendas allow those in power to 'get things done'" (Graaf and Wal, p.625). An example is the Dutch Minister of Finance acquiring ABN AMRO Bank in 2008 without informing Parliament. This was against the law, but time was short due to the impending credit crisis. In order to effectively govern, the minister broke the law. His governance was effective, but not ethical. If he had been ethical, he would not have been effective (Graaf et al. 2016). This same clash of values appears in matters of surveillance. What is more important – effective surveillance technology that obtains the needed information to thwart criminals and inform policy makers or protecting the privacy of unconcerned civilians to the extent that needed information is potentially missed? And what takes priority – keeping the cost of such technology low or spending more taxpayer money to invest in technology that better protects privacy and/or more effectively supplies needed information? The value conflict reveals a further dimension – while policy makers struggle with the bureaucratic aspect of satisfying diverse constituents, they and society as a whole also struggle with these three elements as conflicting values and which to give priority.

These different theories, along with Wilson's work, shed light on why there is so little evaluation of the effectiveness of surveillance technology, why any evaluation that is done is performed almost exclusively by the intelligence agencies themselves, why legislative bodies are unable to give one goal priority, and why these agencies are caught in this trilemma of being called on to equally address these three elements of effectiveness, proportionality and cost and yet are forced to perform a trade-off. The structure and nature of bureaucracy is a large contributor, while the trilemma theory states that achieving all of three conflicting goals is impossible, and the trade-off concept argues that public governance must be either ethical or effective, but cannot simultaneously be both. These support what Wilson identifies as an inherent problem with governments and governing – desirable, yet conflicting values and the government's duty to meet them all despite the fact that they are incompatible. This suggests that certain elements of effectiveness evaluation will not change since surveillance cannot extricate itself from bureaucracy.

Implications for the surveillance debate

This article's findings impact the broader surveillance discussion by revealing that government behavior related to surveillance effectiveness is not something that can just be changed by a new law or new proscriptives. Realizing that there are bureaucratic constraints that obstruct evaluating overall effectiveness, influcences the discussion. The surveillance discussion should take this factor into account, thinking in innovative ways to include effectiveness evaluation, while at the same time realizing there will always be bureaucratic elements to contend with (e.g. elected officials who report to their constituencies, intelligence agencies with primary and contextual goals, the unobservable outcomes inherent in intelligence work). More research is needed that addresses this overarching question of evaluating the effectiveness of surveillance technology in intelligence agencies. Future work could explore how effectiveness can realistically be assessed alongside considerations of privacy issues and cost. Possibilities could include investigating how intermediaries, such as private companies or individuals, might be employed as independent third parties to assess effectiveness. An example is the case of David Anderson Q.C. in the U.K. – Independent Reviewer of Terrorism Legislation – who was asked by the Home Secretary to evaluate and report to the prime minister on the effectiveness of four bulk powers (Anderson 2016). The Reviewer is completely independent from the government, and has a high-level security clearance, providing access to classified information. As the title indicates, the U.K.

Independent Reviewer's role is to review and inform on terrorism legislation, not surveillance technology per say. However, in this particular report Anderson was asked to review the effectiveness of surveillance programs. A powerful element of Anderson's report is that it is designed to inform the public. It is not a classified document with only portions made public. A drawback it that it explicitly excludes considerations of proportionality. Future work could consider such an independent type review of effectiveness that includes proportionality and cost considerations. Moving the issue of effectiveness evaluation outside the government to an independent third party may help free it from some of the encumbering bureaucratic elements.

Wilson's conclusions regarding effectiveness being difficult to evaluate are, of course, applicable to all procedural government institutions, not just intelligence agencies. Wilson himself does not specifically discuss intelligence agencies. This paper has applied his discussion of bureaucracy and effectiveness directly to intelligence agencies. The realm of intelligence is often perceived as a black hole by the public and merits having a direct link made with Wilson's discussions of bureaucracy and effectiveness.

6.6 Conclusion

Evaluating the effectiveness of surveillance technology in intelligence work appears to be nearly non-existent both by the intelligence agencies themselves and by their oversight bodies. Why is this so? And why do intelligence agencies have apparently conflicting goals, and intelligence officers speak of proportionality as a goal they themselves do not address? This article began with a hypothesis that the concept of bureaucracy could shed light on these findings. To test this hypothesis, it examined several observations about bureaucracy in James Q. Wilson's book *Bureaucracy*. Applying these concepts to this larger study's previous research findings revealed that intelligence agencies have conflicting goals because the legislative bodies who direct them cannot decide which goal, among a multitude, to give priority. Effectiveness is minimally treated because intelligence agencies are procedural organizations in which outcomes are difficult to evaluate, and because both workers and managers tend toward what is easier to perform and to measure (e.g. cost and legality rather than effectiveness). Intelligence agencies being procedural organizations also means that they are difficult to oversee. Oversight bodies put rules and procedures in place regarding how to perform tasks, but they do not help in evaluating how well the job has been done. Intelligence oversight, therefore, has a fundamental problem of not assessing the effectiveness of a completed task. Lastly, politics in a democratic government play a role. Parliamentary bodies represent constituencies with various and opposing views and therefore cannot arrive at a consensus regarding which goal should be given priority. Further, a democracy must carry out certain tasks, including conducting surveillance. The public demands that it does so effectively, at low cost, while protecting privacy. Being beholden to meet these conflicting demands is a result of operating in a democracy.

The current status of evaluation and lack of evaluation of the effectiveness of surveillance technology will likely continue as the status quo. Democratic bureaucracy will continue to be in deadlock over which goals take priority, oversight bodies will persist in avoiding the difficult task of assessing effectiveness, as well as in avoiding making trade-offs, workers and managers of intelligence agencies will continue tending towards the more easily measured elements of cost and legality, and security will remain as a primary goal of intelligence agencies while proportionality remains contextual. The necessity of making a trade-off and of not being able to simultaneously address all three elements of overall effectiveness will remain. These are the fetters of bureaucracy.

This is not to say that all is hopeless static. There are perhaps certain adjustments or changes that could be made to improve or make possible evaluations of effectiveness, which future research could explore. However, it should be done knowing that intelligence agencies operate in a democracy, and that, therefore, certain elements of bureaucracy are at play, which cannot be altered.

Chapter 7
Conclusion

This dissertation has brought together two themes: the surveillance debate, which often turns around questions of privacy and proportionality and if this must be exchanged for greater security; and the effectiveness of surveillance technology. These two themes were joined in the research question, *How is the effectiveness of surveillance technology evaluated in intelligence work?*

This concluding chapter will summarize the findings of previous chapters, linking them to the two themes of the surveillance debate and the effectiveness of surveillance technology. This will be followed by reflections on additional actors in the surveillance discussion that were not included in this research, as well as on the utility of frameworks – in this case, frameworks of effectiveness evaluation. The chapter finishes with the scientific and societal implications of the dissertation's findings, and considerations for future research.

7.1 Summary of Findings

Chapter 2 was written in the aftermath of the Snowden leaks, which were, in fact, the driver behind this chapter. The chapter examined various NSA surveillance technologies discussed in the press and classified them on a scale of bulk collection to targeted surveillance. In the months following Snowden's release of classified documents, media reports were full of accusations of mass surveillance by the NSA and its U.K. counterpart, GCHQ. The information obtainable from the leaked documents was piecemeal, and the media reports often appeared to jump to conclusions without much in-depth inquiry and consideration of what the documents might mean. Likewise, there were many different surveillance programs, resulting in a confusing array of technologies and an impression that every technology was sucking up all the information of every innocent citizen.

The purpose of this chapter was to create some sense out of all the hype and to categorize the surveillance programs according to whether they performed targeted surveillance or bulk collection or something in between. In order to talk about effectiveness, one must first know what kind of technologies are in question. Understanding how they operate and the general security purpose they serve is a necessary first step. The technologies discussed here is a sampling of the kinds of surveillance technologies intelligence officials, oversight bodies, and the public are talking about in the following chapters of this thesis (Chapters 3-5).

Chapter 2 found that certain NSA technologies classified as targeted, while others categorized as mass surveillance/bulk collection, and still others fell mid-spectrum. Further, the position of certain technologies on the scale varied according to how they were used in a given case. Subsequent research revealed additional nuances to these classifications, such as how one defines "mass" surveillance, and if a database is a surveillance program if the information has previously been collected by another technology. This chapter treated the database XKEYSCORE as a surveillance technology because it was often mentioned as such in reporting of the time, in discussions of mass surveillance. However, it could easily be argued that this system should not even be discussed as surveillance technology as it performs no actual collection itself. Rather, it processes data that has been collected by other technologies.

Later research (see next chapter) also shed light on the gross disparity between privacy advocates on one side accusing intelligence agencies of conducting mass surveillance through the use of such technologies, and intelligence officials on the other side emphatically denying such claims. This chapter's classification of technology should bear these nuances and differences in mind – i.e. the category of "bulk collection" is what privacy advocates have called "mass surveillance," and as previously mentioned, databases could be left out of this discussion since they are arguably not surveillance technology.

In principle, the research in this dissertation applies to all kinds of surveillance technology. In practice, however, most of the data collected and analyzed is related to the kinds of technology discussed in this chapter. Beginning with a classification of technologies provides a knowledge of the types of technologies in question when discussing effectiveness. It also allows for later research to evaluate the technology's effectiveness according to its position on the scale.

Chapter 3 examined what American and British intelligence officials have to say about the effectiveness of surveillance technology. It analyzed public material – speeches, articles, etc. – as well as drawing on interviews conducted by the author, to determine how intelligence practitioners themselves consider and evaluate effectiveness. This chapter coined terms of "strict effectiveness" and "overall effectiveness." The former refers to whether or not a given surveillance technology achieves the desired security goal. However, this is not the sole criteria to determine whether or not to employ the technology. Effectiveness is never determined in a vacuum. The cost of the technology and questions of proportionality play into whether it is effective "overall" and therefore deployable. Overall effectiveness, thus, includes considerations of strict effectiveness, cost, and proportionality.

Intelligence officials were found to minimally treat the question of strict effectiveness. It was rarely addressed in the material analyzed for this study. Officials stated that evaluating effectiveness is extremely difficult, if not impossible, especially when it concerns strategic intelligence (as opposed to tactical intelligence). Measures of effectiveness used by officials were identified by teasing them out of practitioners' discussions of surveillance programs. The following seven measures were identified: thwarted attacks, lives saved, criminal organizations destroyed, output, context, support, and informed policy maker. Cost, although rarely discussed, was confirmed through interviews to be a significant factor in determinations of surveillance technology employment. Much of practitioners' statements turned around questions of proportionality and privacy, which is not at all surprising given that most were made in the years following the Snowden leaks.

Key findings of this chapter illuminated the dilemma mentioned above between privacy advocates and intelligence officials' opposing statements regarding mass surveillance. Privacy advocates argue that huge amounts of data are collected, including the data of innocent, unconcerned people. Because filters are set wide on deep packet inspection type technology, this is a sort-of indiscriminate collection of data, sweeping of mass amounts of (unnecessary) information. Intelligence officials, on the other hand, call this bulk collection, and argue that it is not mass surveillance because 1) the definition of mass surveillance is collecting *everyone's* data, which this does not do, and that while it does collect a lot of data it is a very small fraction of the amount of data flowing over the internet – in the case of the NSA, 1.6% of internet traffic. 2) This captured data is not seen by human eyes, but is collected by computers based on algorithms. Analysts then search this data according to certain, specified criteria (in the case of GCHQ laid out in search warrants). Only data which matches these criteria is shown to the analyst. Thus, human eyes see only a very small percentage of data, which must match specific criteria – in the case of the NSA only 0.00004% of the world's internet traffic is seen by analysts.

These findings shed light on how a critical stakeholder – intelligence practitioners – in the surveillance debate considers questions of overall effectiveness. Equipped with this understanding, other stakeholders can better engage with intelligence officials, resulting in more fruitful discussion.

Chapter 4 turned to oversight bodies and how they evaluate the effectiveness of surveillance technology. Oversight bodies were found to avoid evaluating strict effectiveness. Only two instances were found in which oversight bodies specifically addressed effectiveness. Once was in response to an internal complaint; the second was in response to public outcry in the wake of the Snowden leaks – the PCLOB assessed the effectiveness of two NSA surveillance programs in order to address proportionality. Aside from these instances, oversight mechanisms push the intelligence bodies themselves to evaluate effectiveness. Although oversight bodies themselves rarely evaluate effectiveness, they were found to value the same measures of effectiveness as intelligence officials. This is not surprising given that oversight bodies depend on intelligence practitioners to assess effectiveness, as well as to convey how to measure effectiveness. Practitioners are closest to the technology, possessing the necessary expertise and understanding of its (potential) value. This, as well as an (inevitable) reliance of oversight bodies on intelligence agencies to supply the necessary documents and testimony for investigations, results in an interdependence between the two entities. This does not mean that oversight bodies do not find fault with intelligence agencies' actions, but it does help explain why certain privacy advocates distrust oversight.

A significant finding was that oversight avoids addressing effectiveness, cost, and proportionality simultaneously. It does not like engaging in trade-offs and thus addresses only one or at most two elements of overall effectiveness at once. It thus avoids entering into the impossible trilemma of successfully evaluating all three simultaneously. This is reflected in the fact that oversight bodies are created to address only one issue: the PCLOB, as the name suggests, deals with matters of privacy; the Interception of Communications and Intelligence Services Commissioners both report on legal compliance; other bodies' mandate is to oversee spending.

This chapter began to develop the trilemma concept. This concept exists in macroeconomics – three conflicting, yet desirable goals which cannot all be achieved simultaneously. Here it was applied to the security realm. This was found to be an important concept in this thesis, as it explains the dilemma, or rather, trilemma, that the intelligence community is in related to surveillance technology. In principle, no stakeholder is against effective, cost-efficient, proportionate surveillance technology. However, in reality it is not possible to have all three at the same time. Thus, parties must choose two out of the three. Which two are chosen depends on which stakeholder is doing the choosing, as well as on things like the political climate, recent events, etc. Other stakeholders may not agree with this prioritization and cry foul play.

A final stakeholder was analyzed in **Chapter 5** – the public. The study explored the public's perceptions of surveillance technology effectiveness, as well as possible correlations between perceptions of effectiveness and privacy, and between effectiveness and acceptable cost of the technology. Conducting surveys of Dutch university students and their parents, this research found that the public does not engage in trade-offs in this debate. That is, it does not accept less effective surveillance technology for more privacy and vice versa. Rather, it expects the government to employ effective technology, at a reasonable price, while protecting citizens' privacy.

Significantly, notable differences were found between the student and parent generations. This appeared in correlation analysis between effectiveness and privacy. Parents were found to be divided into groups of trusting or concerned – either they were trusting of intelligence bodies and ISPs' handling of their data (i.e. they did not feel their privacy was breached) while simultaneously they viewed surveillance technology as effective, or they were concerned that their privacy was being invaded while at the same time they considered the technology to be ineffective. In other words, participants did not consider that the technology was effective, but they were having to give up their privacy – i.e. make a trade-off – rather, the technology was both privacy-invasive and ineffective. However, this correlation did not hold true for students. Students were not either fully trusting or concerned and showed no correlation between effectiveness and privacy. This may reflect that students are still forming their views on security matters, while their parents' views are already solidified and perhaps politicized, as questions around security often become highly political.

Other significant findings were that students in the SSJ minor showed lower expectations of effectiveness on the part of the government than students not in the minor. Finally, both participants who judged NSA surveillance technology to be effective and those who considered it ineffective gave the same top explanation as to why – too much data collected.

The public is an important stakeholder as its views on the effectiveness of surveillance technology potentially influence its acceptance of the technology, and ultimately laws and policy governing the technology's use. Government officials often speak in terms of trade-off. Knowing that the public does not, in fact, think in terms of trade-off related to any of these elements of effectiveness, cost, and proportionality/privacy is crucial for these parties to engage in meaningful dialogue.

Chapter 6 investigated the underlying reasons for the lack of effectiveness evaluation. Starting from the hypothesis that the institution of bureaucracy may explain this lack, this chapter applied concepts from James Q. Wilson's book *Bureaucracy* to previous chapters' findings. The resulting conclusion was that effectiveness evaluation is, in fact, fettered by bureaucracy. Bureaucracy, a necessary institution in democracies, issues in: 1) intelligence agencies having multiple, conflicting goals because their governing legislative bodies cannot decide which goal to give priority; 2) intelligence bodies being procedural organizations with unobservable outcomes, which means that success is difficult to evaluate and the agencies are difficult to oversee; 3) intelligence employees and managers gravitating towards the easier, more measurable benchmarks of cost and legality, rather than effectiveness. Oversight bodies require rules and procedures to follow, but do not evaluate the success or effectiveness of completed tasks. And while intelligence agencies are not political bodies, they are affected by the politics of democracy as seen with the aforementioned legislative bodies, which do not prioritize conflicting goals, and with the inherent conflict of being beholden to satisfy conflicting demands made by the public.

Additional theories buttressed and gave further dimension to these findings. Oversight bodies are compartmentalized to address specific issues and distance themselves from trade-offs. Governance literature discusses value conflict, arguing that effective and ethical governance clash. Applied to the security realm, not only do policy makers struggle with the bureaucratic aspect of satisfying the conflicting demands of their constituents, but they and society as a whole struggle with the conflicting values of effective surveillance, privacy, and cost, and which to prioritize.

This chapter ends with a proposal to explore engaging independent third parties to conduct overall effectiveness reviews, which would examine the strict effectiveness of given surveillance technologies, while at the same time incorporating considerations of cost and proportionality. This would not be a solve-all-problems solution, but it may be a way to free effectiveness from some of the fetters of bureaucracy. Any proposal to introduce or improve evaluations of effectiveness should be done knowing that bureaucratic constraints are at play, which cannot be eliminated.

In summary, the initial response to the main research question of how the effectiveness of surveillance technology is evaluated in intelligence work, is that it rarely is. Digging deeper, however, reveals that stakeholders do hold measures by which they inherently evaluate effectiveness. Intelligence practitioners and oversight bodies hold very similar measures of effectiveness. They also both live in the reality of the security trilemma, never assessing effectiveness, cost, and proportionality simultaneously, and consequently often speak in terms of trade-off. The public, on the other hand, carries an opposing perspective of not engaging in a trade-off and expecting all three elements to be delivered together. As far as its own evaluation of effectiveness, the public has perhaps entirely different measures – measures based on political party or leanings, on trust in government, on more absolute positions regarding privacy and security (e.g. being concerned about privacy regardless of the institution), or on some unknown value.

Perhaps no stakeholder considers the role played by bureaucratic constraints and the realities imposed by bureaucracy.

7.2 Reflections

7.2.1 Additional Actors

This dissertation examined the views on effectiveness of several stakeholders in the surveillance debate – intelligence practitioners, oversight bodies, and the public. Additional stakeholders that could have been included but were not were, notably, law enforcement, privacy advocates, and private companies. This section reflects on these additional actors.

Law Enforcement

Law enforcement's inclusion as a stakeholder could be debated. The focus of this thesis is on surveillance technology employed in intelligence work. Intelligence work is primarily the work of intelligence agencies, not law enforcement. The purpose of intelligence is to inform – it comes before any action is taken and informs officials and policy makers. Based upon the intelligence they make a decision. Law enforcement's function is to respond to a specific incident or indication of criminal activity. Intelligence agency action could be said to be pre-incident, while law enforcement action is post-incident. These differences and others are summarized in the table below.

Law Enforcement	**Intelligence Agencies**
Court of law – what LE collects has to stand up in a court of law	Foreign policy – intelligence collection influences and directs foreign policy decisions
Investigate to prosecute – must always have a mind to holding information in evidential form	Investigate to collect info, inform heads of state, for policy decisions
Domestic	International
Transparency – has to show operating according to law	Secrecy
Tactical – almost exclusively	Strategic & tactical

Table 7.1. Differences between law enforcement and intelligence agencies

However, current security threats, particularly terrorism, have resulted in a shift in which law enforcement is engaged more in intelligence. At times, law enforcement and intelligence agencies work together to surveille a suspected terrorist group and collect evidence against them. The obvious goal with counterterrorism work is to use intelligence to preemptively stop a terrorist group from acting. The place of intersect for law enforcement and intelligence agencies are these pre-emptive tactical operations. Two interviews conducted for this thesis, focusing on a British counterterrorism case – Operation Pathway – highlight this intersection. This case was conducted conjointly with law enforcement and MI5, Britain's domestic secret service.

Initial investigation into these types of law enforcement investigations suggests that effectiveness evaluation might be more present than it is with intelligence agencies. Interviews were conducted with a British Detective Superintendent who had been part of a counterterrorism unit. The results of these interviews indicated that British law enforcement monitors and reviews the operational benefits of its surveillance technology. This includes assessing the technology during deployment to determine if it is achieving the desired goal, seeking advice and support from other police forces and security service representatives when particular difficulties are encountered, and quarterly reports on positive and negative experiences with specific technologies. Some of these, however, are more focused on technical issues rather than effectiveness. Law enforcement evaluation of effectiveness could be explored further in future research, both in intelligence work and criminal investigative work.

Privacy Advocates

This dissertation intentionally decided not to cover the views of privacy advocates related to the effectiveness of surveillance technology, primarily because privacy advocates are arguably the loudest voice when it comes to discussing surveillance technology. Their views are well-voiced and well-known. The views of government institutions who deal firsthand with surveillance technology – intelligence agencies and oversight bodies – are much less known, making this focus vastly more interesting to study. Privacy advocates have an obvious focus on privacy rather than effectiveness, arguably making a study on their views of effectiveness less pertinent for this research.

Private Companies

Private companies' evaluation of effectiveness is worthy of an entire thesis. No less so because surveillance technology now includes not only traditional forms such as CCTV, but monitoring online activity through the use of technology such as deep packet inspection. How these companies evaluate the effectiveness of their surveillance technology and balance this against protecting privacy and business costs is a subject that should be investigated. Private companies were not included in this dissertation due to the vastness of this topic, but addressing this subject (including calling it surveillance, which business avoids doing) should be a study for future research.

7.2.2 The Dilemma of Frameworks

SURVEILLE & MOEST models

As with many PhDs, this one has taken several turns over the course of its existence. It began with an intention to create a prescriptive model of how to measure the effectiveness of surveillance technology. This is, indeed, one approach to studying the subject of effectiveness. One could identify certain measures against which effectiveness should be assessed, and even develop a weighting system, with certain measures given more importance than others. Two models to this effect have been developed in connection to this PhD. The first was within EU FP7 SURVEILLE project, of which this PhD was a part. The second was as part of the PhD itself. The dilemma of this approach is that it does not address the underlying *why* of the absence of effectiveness evaluation. If stakeholders have goals other than effectiveness, or if bureaucratic constraints hinder effectiveness assessment, then introducing models and frameworks will not solve the problem. This dilemma is evident with the SURVEILLE model.

The SURVEILLE project developed a matrix to score surveillance technology according to usability, ethics, and human rights. These sections were divided into categories which each received a score. For usability the categories were effectiveness, cost, and privacy-by-design. The intention was that particular technologies would be scored in specific situations across these three subjects of usability, ethics, and human rights with points assigned to each category. The resulting overall score would indicate whether or not the technology should be used in the given scenario. For example, in a hypothetical drugs and firearm investigation, AIS ship location detection scored high across all three sections, while a bug planted in a target's home received a low ethics and human rights score and a high usability score resulting in an overall low score (Guelke et al. July 2013).

The project envisioned a potentially broad application of the matrix – national and European policy makers, technology developers, law enforcement (application in the intelligence realm was not mentioned). The purpose of the matrix, however, was to measure the technology's use in a given circumstance against an external fundamental rights assessment, rather than against the legal basis for using the technology. In other words, it was not to assess whether or not the law enforcement officer acted according to the law in deploying the technology, but to assess the law itself against a human rights and ethics assessment, and whether the use of technology was necessary and proportionate based on an external human rights evaluation (Cayford et al. 2014).

The SURVEILLE matrix is a good example of the futility of such a model if it does not adequately take into the account the realities of the situation – the needs and/or goals of the stakeholders, bureaucratic constraints, etc. Although the leaders of the SURVEILLE project envisioned that law enforcement officials could potentially use the matrix, several factors render this unlikely. Law enforcement officials who were part of the SURVEILLE project expressed dissatisfaction with the matrix – they disagreed with the assessments of necessity, proportionality, and justification being made not according the existing law, but according to an external human rights assessment. Indeed, law enforcement conducts surveillance according to what is legally proscribed. It is unlikely that it would undertake an external human rights assessment in addition to all its other work, particularly when its operations already strictly adhere to the law. David Anderson in his Report of the Investigatory Powers Review also questions the lack of consideration in the SURVEILLE matrix for oversight and regulation to reduce ethical risk (Anderson 2015). This is, effectively, a noticeable lack. The SURVEILLE matrix lacks taking account of realities – expressed concerns of the matrix not accounting for the necessity and proportionality judgments officers must make, time and resource to devote to yet another assessment/ requirement that does not account for existing law, oversight, and regulation. This is, of course, just one application for the matrix. The matrix could be more fit for other envisioned applications of informing policy makers in drafting law.

The second model was developed exclusively within this PhD research – Measures of Effectiveness for Surveillance Technology (MOEST). The purpose of this model was to provide a structured starting point for eventual effectiveness evaluations. Given that surveillance technologies differ and that each use is a unique situation, this model was envisioned to serve as a guideline, which could be adapted to individual applications. MOEST is based on Sproles' work on measures of effectiveness (1999). Sproles asserts that in order to determine specific measures of effectiveness, stated goals are needed. In other words, before a technology can be evaluated for its effectiveness, the goal of the surveillance technology's deployment in a specific situation must be determined. The technology's effectiveness is then measured against this goal. Practitioners are those who determine the goals of surveillance technology; in intelligence work these goals are typically secret. Furthermore, each deployment of surveillance technology requires specific measures of effectiveness tailored to that task. It is, therefore, not possible to create a method for measuring effectiveness whose values are applicable to all uses of surveillance technology. Thus, MOEST was developed as a method of general measures, intended to serve as a guide from which stakeholders could adapt measures and assign values specifically tailored to each situation.

MOEST is divided into three categories, which provide different angles for evaluating the surveillance technology: quantification, knowledge gained, and support. These measures include a combination of more easily quantifiable means of "counting" incidents and less easily quantifiable measures, such as knowledge acquired. This is consistent with the purposes of intelligence. The collecting of intelligence and the goal of its collection serves a broad purpose of informing heads of state, which is difficult to quantify. Further, because bits of information typically come from a multitude of sources, together they inform the whole picture, making it more difficult to attribute a success to one kind of technology or another. These measures try to take these qualities of intelligence gathering and purpose into account, and create a balanced spectrum that includes these different considerations. The MOEST method appears as follows:

1) Quantification

- Number of reports generated containing information collected by the surveillance technology
- Number of plots/ attacks thwarted
- Number of arrests leading to prosecution/ captures/ kills

2) Knowledge gained

- Identification & location of suspects
- Strategy, hierarchy, modes of operation of an organization

3) Support

- Corroborating information in other investigations
- Relationship to and support of other surveillance technology/ programs
- Dissemination of information to other agencies
- Warnings of new and ongoing threats

Figure 7.1. The MOEST method

Such a model could prove useful to creating some structure for evaluating effectiveness. Further, it is broad enough that it could be adapted for tactical versus strategic intelligence operations, for example. Thus, this PhD could have continued down the prescriptive road of further developing and discussing this model, introducing a weighting system of the measures, etc.

The dilemma of such an approach is that while, in theory, such a model may be useful, if it is never used by those for whom it is designed, its existence becomes pointless. While there appears to be a lack of models of effectiveness evaluation in the intelligence realm, more fundamentally there is a lack of assessing effectiveness period. It would be naïve to assume that this lack exists because intelligence practitioners or oversight bodies do not have an interest in effective technology. Therefore, introducing nifty models will not solve the problem. If stakeholders do not evaluate effectiveness because they have other goals or because bureaucratic constraints limit their ability to do so, introducing frameworks is not going to solve the problem. Finding and examining the empirical evidence behind this absence began to make more sense for this research. With this starting point, future research could perhaps take on a prescriptive angle that would be more effective, taking into consideration the empirical findings regarding why the lack of evaluation exists.

7.2.3 The issue of trust

Trust is an underlying and crucial theme when it comes to surveillance and intelligence agencies. It influences views of effectiveness and perceptions of privacy invasion. The issue of trust surfaced in the research on oversight bodies and the public. Because intelligence agencies operate in secret, the public seeks and needs some reassurance that it can trust their actions to be within the law. Therefore, oversight bodies are established. But then, how can these oversight bodies be trusted? This dissertation revealed that governments strive to solve this problem, at least in part, through transparency. British oversight bodies, in particular, seem to recognize this connection. Their reports often underline that, in order to gain public trust, their oversight must be as transparent as possible. This is one of two facets of transparency and oversight that this study identified. The second is the transparency of the intelligence agencies vis-à-vis the oversight bodies. The more the intelligence agency trusts the oversight body, the more likely they are to provide the requested documents in a timely manner. And conversely, the better they cooperate, the more likely the oversight body is to trust them.

The factor of human judgment in proportionality decisions discussed in Chapter 4 also comes back to trust. Those who trust the oversight mechanism and its members and/ or the intelligence practitioner, trust their judgment. Those who are fundamentally distrustful of the intelligence community distrust practitioners' judgments, as well as seemingly all (e.g. oversight bodies) who come to a conclusion in support of intelligence agencies' judgments of proportionality. David Anderson, in his independent review of U.K. terrorism legislation came to this ultimate conclusion, that the question of trust is a core issue (the title of his report – "A Question of Trust" – reflects this). His recommendations were formulated with the need to promote trust in mind.

It could be said that the core issue is not whether or not the surveillance technology at hand is effective or even whether the surveillance actions are proportionate, but whether the public and other stakeholders trust the institution in question. Studies have found that the public is divided into two groups of either trusting or concerned (Pavone and Degli Esposti 2010, Van den Broek et al. 2017). This PhD found that this division is age-related, pertaining to adults, but not to university-aged students. This could be related to politics – one's view of intelligence services may be more affected by the political party and/or head of state in power, rather than the agency itself. The head of state is perceived as responsible for what the intelligence services do, and therefore if one holds a favorable view of the president or prime minister, this extends to the intelligence services. Politics may not be the sole explanation, however. The research for Chapter 5 found that people (adults) are trusting or not regardless of the institution involved. One of the types of institutions respondents were questioned about was internet service providers, and the degree of trust in these private institutions paralleled the degree of trust towards government agencies, domestic and foreign. This may suggest that other factors are at play.

A paradoxical relationship exists between trust and bureaucracy. Citizens of democratic societies strive for a society of equal opportunity and are often distrustful of people in power. Therefore, they create institutions to keep those in power in check, to institute rules of justice and fair play, to hold government officials accountable for their actions. And yet these controls do not necessarily result in trust, but in a distrust of the institutions themselves. Democracy cannot exist without bureaucracy, and yet citizens become distrustful of the very bureaucracy they have instituted.

7.3 Implications

7.3.1 Scientific Insights

As seen in the literature reviews in the previous chapters, surveillance literature tends to circle around questions of privacy. Where effectiveness has been treated, it is in the context of counterterrorism programs and cost-benefit analysis. And its examination in relation to surveillance technology is largely limited to CCTV. The effectiveness of the surveillance technology in intelligence agencies in meeting a given security goal has been a neglected topic. This dissertation contributes to filling this gap in the literature through five principle insights:

Non-existent effectiveness evaluation. Intelligence agencies and their oversight bodies rarely evaluate the effectiveness of surveillance technology. Effectiveness appears to be valued primarily as it relates to cost, rather than as an element in and of itself. Agencies and policymakers seek out technologies that give good results vis-à-vis the amount of funds being spent. Budgets are fixed and limited, so these cost-benefit analyses comparing different technologies are desirable. Evaluating the actual effectiveness, independent of cost, is difficult for oversight bodies to perform as they lack familiarity with the technology. Analyzing statements of intelligence officials revealed that this kind of evaluation is fraught with difficulty – where is the threshold of effectiveness (i.e. what percentage of the time must it yield results to be considered effective), should it be evaluated on its own or in conjunction with other technologies and programs, where does this group of technologies end, how can effectiveness be quantified, etc.? This is but part of the "why" behind the lack of effectiveness evaluation, which is another contribution addressed below, but already determining that this lack exists is an important insight. Scientists could use this insight to investigate whether this lack exists in other arenas – law enforcement, intelligence services in other countries, the private sector, environmental programs, etc. It raises the question of how many programs and technologies in all types of sectors are put in place without evaluating their actual effectiveness.

Measures of effectiveness. Although effectiveness evaluations were found to be lacking, measures that officials inherently use or value were determined through analysis of public documents, speeches, and interviews. Technology elicited as successful points to measures of effectiveness valued by intelligence practitioners and oversight bodies. Understanding their frame of reference is important both for dialogue – discussing privacy issues, surveillance policy and law, etc. – and for any future development of models of effectiveness evaluation. Scientific research in both these areas of privacy and policy matters vis-à-vis surveillance and developing effectiveness models would benefit from bearing these measures in mind as the research is conducted.

The security trilemma. Government agencies seek to fulfill three desirable yet conflicting goals – employ effective surveillance technology, at a reasonable cost, while protecting privacy. Because it is impossible to accomplish all three simultaneously, they must give up at least one while striving for the remaining two. This is reflected particularly with oversight bodies who never address all three of these elements at a given time. This impossible trilemma is avoided primarily by oversight bodies being created to deal with only one of these goals. If a body has the mandate to address all three, it avoids the trilemma by issuing reports that evaluate or address a maximum of two goals at once. This reflects a reality of it being impossible to simultaneously have maximum effectiveness and maximum privacy protection at the lowest cost. The most effective surveillance technology will inevitably have a high price tag and collect a maximum amount of information. Likewise, maximum privacy protection will carry a high likelihood of missing pertinent information, and an inexpensive technology will likely lack performance, rendering it impossible to be effective. Perhaps a highly effective technology with good privacy protection measures is possible, but it would come with a high price tag, likely putting it outside of budget possibilities. This security trilemma is a harsh reality, not something that can be brushed aside or ignored in the surveillance debate. Acknowledging this reality in this larger debate is a necessary part of future research in this field, whether it be focused on privacy or legal matters, the technology itself, or the institutions involved. The security trilemma is equally applicable to studies in sectors of "non-traditional" surveillance, such as gas/electricity smart metering, health care monitoring systems, and collection of online data for commercial purposes.

Fetters of bureaucracy. The source of the lack of effectiveness evaluation in intelligence agencies is found in the institution of bureaucracy. Bureaucracy is a necessary part of democratic government, and yet the requirements put upon the government by the people themselves hinder the assessment of effectiveness. The knowledge that these fetters exist and cannot be done away with is crucial to the surveillance debate and to potential policy implications. This allows for realistic approaches and proposals to introducing evaluations of effectiveness. More broadly speaking, research investigating questions of governance related to surveillance technology, intelligence and oversight reforms, and government reforms in general would benefit from considering the impact of bureaucratic constraints.

Impasse of public demands vs. impossible trilemma. As seen above, the government finds itself in an impossible trilemma of trying to simultaneously deliver three conflicting goals (effective, cost-efficient, proportionate), and therefore, being forced to choose two out of the three at any given time. Juxtaposed against this is the public who demands that all three be delivered simultaneously. These two opposing perspectives lead to an impasse. It is the reason government officials speak of trade-off; it is the reason recent studies (Degli Esposti et al. 2017, Van den Broek et al. 2017) have concluded that the notion of trade-off is misaimed. This literature perhaps presumes that government officials address the public in terms of trade-off because they think the public engages in these terms, that it views these subjects of surveillance in a weighting/balancing perspective. Whether or not this is the view officials hold, their trade-off dialogue stems from a deeper reality of the afore-mentioned security trilemma, not mere talk to match what they presume members of the public are thinking. Future surveillance studies would do well to integrate this impasse into their analysis and discussions.

7.3.2 Social Implications

Out of the Impasse

Effectiveness is not only the intersect where technology and policy meet, but may also be the road leading out of the impasse between the public not engaging in trade-offs and the government having no other way (trilemma) but to perform trade-offs. As Chapter 6 proposed, an overall effectiveness evaluation assessing effectiveness, cost, and proportionality, performed by an independent third party could be an effective way out of this impasse. The reviewer would be independent from the government, but highly qualified and possessing the necessary security clearances to conduct such a review. While having access to classified information in order to conduct the review, sensitive information would contribute to the report's conclusions, but would not be in the report itself. The effectiveness report would be fully public, as in the cases of David Anderson's *Report of Bulk Powers Review* in the U.K., and the PCLOB's reports on NSA programs under Sections 215 and 702.

This would provide a transparent way to evaluate effectiveness. As such, it avoids the never-ending argument of privacy versus security and instead addresses another important yet neglected element of the surveillance debate – effectiveness. It is in nobody's interest to have ineffective surveillance technology used in intelligence work. For a moment at least, questions of how much data should be collected, etc. are put aside as the question of *is the technology achieving the said security goal* is investigated. This is determined by the reviewer as it was in Anderson's report. The challenges of evaluating effectiveness are thus made evident to the public, and simultaneously, the measures used to evaluate effectiveness, the steps of evaluation, and the ultimate conclusion are known. The report would then go further to evaluate the technology's proportionality and cost. In this way the realities of the security trilemma are laid out, and the public is brought into a dialogue that more meaningfully addresses the root issues of the surveillance debate.

Obviously, not everyone will agree with the conclusions of such an evaluation, and the purpose of such an assessment is not to put the debate to rest. Surveillance and its boundaries is a heated topic and democracies should debate its implementations. But an evaluation such as this serves two purposes: 1) it breaks out of the never-ending tug-of-war of privacy vs. security, and provides for a more meaningful debate by bringing in neglected elements, as well as hopefully providing a less charged starting point of focusing on effectiveness. The brings the debate beyond the level of making general statements, such as, “we need this technology to secure our nation,” to an actual evaluation of, is this technology effective, and then is it proportional, etc. 2) It introduces more transparency to intelligence agencies. Some will argue that the required secrecy around intelligence methods makes such an evaluation impossible. Anderson’s report proves this to be false, as well as the reports by the PCLOB. In both cases enough information was given for the public to be adequately informed, and yet sources and methods were protected. Intelligence agencies recognize that society has entered an era in which trust is not implicitly given by the public, but rather must be earned and maintained by the intelligence services. More transparency is one way to do this.

Policy

While the above is one proposition for effectiveness evaluation, more generally speaking, these findings provide a starting point from which to seek to establish an objectified manner of evaluating the effectiveness of surveillance technology. Currently, there is no established, objectified way of looking at these criteria, which means that they will be politicized. In developing policy related to surveillance technology, effectiveness (and not only matters of privacy) should be recognized and included as an important element, as well as the realities of the security trilemma and the fetters of bureaucracy. The focus should not only be on privacy issues, but also on effectiveness, recognizing and addressing this as equally important as privacy and proportionality.

More effective dialogue

These findings provide depth and content to various stakeholder views, providing the way for more meaningful discussion around these complex issues. Stereotypes tend to exist in the way of government officials who pursue surveillance and security at all costs, throwing privacy to the wind. The public, in turn, is often perceived and portrayed as holding a trade-off mentality. Holding these stereotypes risks not truly hearing and understanding the actual perspectives of these stakeholders. Delving into these perspectives reveals the complexity of the subject.

Dialogue that discusses in more depth what effectiveness means and how it is measured, that brings to the forefront the security trilemma, that discusses trust and the central role this plays, that recognizes and addresses head-on bureaucratic constraints – this is the kind of dialogue needed in the surveillance debate. The political conversation has to shift from saying we do or should deliver effective, cost-efficient, and proportionate surveillance technology, to asking what we are willing to sacrifice. The public must be made aware of the complexities of the subject. Discussing the security trilemma and the fetters of bureaucracy is a way to make the conversation more honest – only two of the three values can be taken in at once and this within bureaucratic constraints.

7.4 Limitations

This dissertation is limited by the ambitious choice to study intelligence agencies. As is well known, these agencies operate in secret. Data for this research has been collected from open sources. Contradictory data, therefore, may exist within the intelligence services. This study has focused on intelligence services and oversight bodies in the U.S. and the U.K. The conclusions drawn might not hold true for intelligence and oversight in other countries. Chapter 5 on the public was based on surveys conducted in the Netherlands. While on the one hand this provides unique insights into Dutch views regarding effectiveness, etc., on the other hand it means that the findings might not hold true for populations of other nationalities. Survey questions also focused on one type of surveillance technology and on private mobile device use in two scenarios. Therefore, focusing on other types of technologies and/or different scenarios might yield varying results.

In analyzing different stakeholders – intelligence officials, oversight bodies, the public – the research question asked varied slightly in each case. For intelligence officials, the focus of the study was what they were saying; for oversight bodies it was how they evaluate; and for the public it was views of effectiveness. These slightly different angles were taken in order to facilitate the research. That is, although one would like to know exactly how intelligence practitioners evaluate effectiveness, this is classified information. Therefore, examining what they *say* about effectiveness provides more solid research footing. From their statements much can be learned about their values and treatment of effectiveness. Oversight bodies, as an entity, "speak" through their documents. Many of these documents are public and drawing conclusions on *how* they evaluate effectiveness is therefore more feasible, which is why this approach was chosen. The public does not evaluate surveillance technology, as such, and regardless of how effective the technology is or is not, it is individuals' views or perceptions of that effectiveness, rather than the actual effectiveness that influence their acceptance of the technology. While these variations in approach may limit the study results in some sense, it did not inhibit the yielding of many rich insights regarding these stakeholders' views of effectiveness.

7.5 Future research

Continuing research on the effectiveness of surveillance technology could take many forms. In line with the proposal of an independent third party to evaluate overall effectiveness, future studies could probe this possibility in more depth. With the findings from this dissertation as a base, other research could take a prescriptive approach, creating models with which to evaluate effectiveness. Of course, additional stakeholders, such as law enforcement and private companies, could be studied regarding their treatment of the effectiveness of surveillance technology. The research on public views of effectiveness should be regarded as just the beginning, and could be taken much farther, extending to other populations and additional technologies and scenarios. Likewise, intelligence services and oversight bodies in other countries could be examined. While non-democracies would obviously have great differences particularly in regards to overall effectiveness, it would be interesting to see how different democracies compare in their treatment of effectiveness, cost, and proportionality. Effectiveness of surveillance technology, as a whole, and particularly as regards intelligence services, is a meagerly researched topic, leaving much room for further studies.

References

Adams, Andrew A. “Facebook Code: SNS Platform Affordances and Privacy.” *Journal of Law, Information & Science* 23, no. 1 (2014).

Adams, Andrew A., Kiyoshi Murata, Yasunori Fukuta, Yohko Orito, and Ana María Lara Palma. “Following Snowden around the World: International Comparison of Attitudes to Snowden’s Revelations about the NSA/GCHQ.” *Journal of Information, Communication and Ethics in Society* 15, no. 3 (August 14, 2017): 311–27.

Aid, Matthew. “Oversight at Last! Senate Intelligence Committee Staff Auditing Every U.S. Intelligence Community Program.” www.matthewaid.com [accessed November 23, 2016].

Aizenman, Joshua, and Hiro Ito. “Trilemma Policy Convergence Patterns and Output Volatility.” *The North American Journal of Economics and Finance* 23, no. 3 (December 2012): 269–85.

Alexander, Yonah, ed. *Counterterrorism Strategies: Successes and Failures of Six Nations*. 1st ed. Washington, D.C: Potomac Books, 2006.

Altheide, David L. *Terrorism and the Politics of Fear*. Lanham, MD: AltaMira Press, 2006.

Ambinder, Marc. “Solving the Mystery of PRISM.” *The Week*, June 7, 2013. http://theweek.com/article/index/245360/solving-the-mystery-of-prism.

Amicelle, Anthony. “Towards a ‘New’ Political Anatomy of Financial Surveillance.” *Security Dialogue* 42, no. 2 (April 2011): 161–78.

Anderson, David. “A Question of Trust – Report of the Investigatory Powers Review,” June 2015. https://terrorismlegislationreviewer.independent.gov.uk/a-question-of-trust-report-of-the-investigatory-powers-review/.

Anderson, David W. K. *Report of the Bulk Powers Review*, 2016. https://nls.ldls.org.uk/welcome.html?ark:/81055/vdc_100035016622.0x000001 [accessed April 24, 2018].

Appelbaum, Jacob. “30c3: To Protect and Infect, Part 2.” presented at the Chaos Communication Congress, Germany, December 2013. http://www.youtube.com/watch?v=b0w36GAyZIA&feature=youtube_gdata_player.

Appelbaum, Jacob, Judith Horchert, and Christian Stöcker. “Shopping for Spy Gear: Catalog Advertises NSA Toolbox.” *Spiegel Online*, December 29, 2013. http://www.spiegel.de/international/world/catalog-reveals-nsa-has-back-doors-for-numerous-devices-a-940994. html.

Armitage, Rachel, Graham Smyth, and Ken Pease. “Burnley CCTV Evaluation.” In *Surveillance of Public Space: CCTV, Street Lighting and Crime Prevention*, 225–50. Monsey, NY: Criminal Justice Press, 1999.

Asghari, Hadi. PhD Candidate, Economics of Cybersecurity, TBM, TU Delft. Author’s Interviews, November 18, 2013 and March 7, 2014. Email correspondence: November 21, 2013.

Asghari, Hadi, Michel van Eeten, and Milton Mueller. “Unravelling the Economic and Political Drivers of Deep Packet Inspection.” Submitted to the GigaNet 7th Annual Symposium. Baku, Azerbaijan, November 5, 2012.

Assistant Attorney General. “Report in Response to FISC Order of July 9, 2009,” August 31, 2009.

Assistant Attorney General, NSA Deputy Director, and General Counsel ODNI. Joint Statement Before the Permanent Select Committee on Intelligence House of Representatives at Hearing on “FISA Amendments Act Reauthorization,” 2011.

———. Joint Statement Before the Senate Select Committee on Intelligence at a Hearing on “FISA Amendments Act Reauthorization,” 2012.

Attorney General. “Minimization Procedures Used by the NSA in Connection with Acquisitions of Foreign Intelligence Information Pursuant to Section 702 of FISA,” September 21, 2016.

———. “Minimization Procedures Used by the NSA in Connection with Acquisitions of Foreign Intelligence Information Pursuant to Section 702 of FISA,” March 24, 2017.

———. “Procedures Used by the NSA for Targeting Non-U.S. Persons Outside the U.S. to Acquire Foreign Intelligence Information Pursuant to Section 702 of FISA,” July 10, 2015.

———. “Procedures Used by the NSA for Targeting Non-U.S. Persons Outside the U.S. to Acquire Foreign Intelligence Information Pursuant to Section 702 of FISA,” March 29, 2017.

Attorney General, and Director of National Intelligence. “Semiannual Assessment of Compliance with Procedures and Guidelines Issued Pursuant to Section 702 of FISA, June 1, 2012 - Nov. 30, 2012,” August 2013.

Ball, James, Luke Harding, and Juliette Garside. “BT and Vodafone among Telecoms Companies Passing Details to GCHQ.” *The Guardian*, August 2, 2013. http://www.theguardian.com/business/2013/aug/02/telecoms-bt-vodafone-cables-gchq.

Barrett, David M. *The CIA & Congress: The Untold Story from Truman to Kennedy*. Lawrence, Kan: University Press of Kansas, 2005.

Bendrath, Ralf, and Milton Mueller. “The End of the Net as We Know It? Deep Packet Inspection and Internet Governance.” *New Media & Society* 13, no. 7 (November 2011): 1142–60.

Bentley, Hannah. “Keeping Secrets: The Church Committee, Covert Action and Nicaragua.” *Columbia Journal of Transnational Law* 25, no. 3 (1987): 601–45.

Bergen, Peter, David Sterman, Emily Schneider, and Bailey Cahall. “Do NSA’s Bulk Surveillance Programs Stop Terrorists?” New America Foundation, January 2014.

Berghel, Hal. “Through the PRISM Darkly.” *Computer* 46, no. 7 (July 2013): 86–90.

Bigo, Didier, Sergio Carrera, Nicholas Hernanz, Julien Jeandesboz, Joanna Parken, Francesco Ragazzi, and Amandine Scherrer. “National Programmes for Mass Surveillance of Personal Data in EU Member States and Their Compatibility with EU Law.” European Parliament, 2013. http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/493032/IPOL-LIBE_ET(2013)493032_EN.pdf [accessed Jan. 10, 2018].

Born, Hans, Ian Leigh, and Aidan Wills. *Making International Intelligence Cooperation Accountable*. Geneva: DCAF - The Geneva Centre for the Democratic Control of Armed Forces, 2015.

Born, Hans, and Aidan Wills, eds. *Overseeing Intelligence Services: A Toolkit*. Geneva: DCAF. Geneva Centre for the Democratic Control of Armed Forces (DCAF), 2012.

Born, Hans, and Marina Caparini, eds. *Democratic Control of Intelligence Services: Containing Rogue Elephants*. Aldershot, England; Burlington, VT: Ashgate, 2007.

Born, Hans, Loch K. Johnson, and Ian Leigh. *Who’s Watching the Spies? Establishing Intelligence Service Accountability*. 1st ed. Washington, DC: Potomac Books, 2005.

Born, Hans, and Ian Leigh. *Making Intelligence Accountable: Legal Standards and Best Practice for Oversight of Intelligence Agencies*. Pub. House of the Parliament of Norway, 2005.

Born, Hans, Ian Leigh, and Aidan Wills, eds. *International Intelligence Cooperation and Accountability*. Studies in Intelligence Series. London; New York: Routledge, 2011.

Brand, Rachel L. “Testimony of Rachel L. Brand,” Member of the Privacy and Civil Liberties Oversight Board before the United States Senate Committee on the Judiciary, 2016.

Brennan, John. “Remarks at the Council on Foreign Relations,” March 11, 2014. on CIA website.

Brown, Dr. Ian, Associate Director of Oxford University’s Cyber Security Centre and Senior Research Fellow at Oxford Internet Institute. Author’s email correspondence, March 25, 2014.

Buch, Nobert, Fei Yin, James Orwell, Dimitrios Makris, and Sergio A. Velastin. “Urban Vehicle Tracking Using a Combined 3D Model Detector and Classifier.” In *Knowledge-Based and Intelligent Information and Engineering Systems*, edited by J.D. Velásquez, S.A. Ríos, R.J. Howlett, and L.C. Jain. KES 2009. Lecture Notes in Computer Science, vol. 5711, 169-176. Springer, Berlin, Heidelberg, 2009.

Caplan, Joel M., Leslie W. Kennedy, and Gohar Petrossian. “Police-Monitored CCTV Cameras in Newark, NJ: A Quasi-Experimental Test of Crime Deterrence.” *Journal of Experimental Criminology* 7, no. 3 (September 2011): 255–74.

Cayford, Michelle, Coen van Gulijk, and Simone Sillem. “SURVEILLE Deliverable 3.8: Report Combining Results of All Effectiveness Research.” Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme, May 29, 2014.

Cayford, Michelle, and Wolter Pieters. “The Effectiveness of Surveillance Technology: What Intelligence Officials Are Saying.” *The Information Society* 34 (2): 88–103, 2018. DOI:10.1080/01972243.2017.1414721.

Cayford, Michelle, Wolter Pieters, and Constant Hijzen. “Plots, Murders, and Money: Oversight Bodies Evaluating the Effectiveness of Surveillance Technology.” *Intelligence and National Security* 33, no. 7 (November 10, 2018): 999–1021. DOI:10.1080/02684527.2018.1487159.

Christensen, Tom, and Martin Lodge. “Accountability, Transparency and Societal Security.” In *The Routledge Handbook to Accountability and Welfare State Reforms in Europe*, edited by Tom Christensen and Per Lægreid, 165–79. London; New York: Routledge, Taylor & Francis Group, 2017.

Churchill, Ward, and Jim Vander Wall. *The COINTELPRO Papers: Documents from the FBI’s Secret Wars against Dissent in the United States*. 2nd ed. South End Press Classics Series, v. 8. Cambridge, MA: South End Press, 2002.

CIA Office of Privacy and Civil Liberties. “Semiannual Report January-June 2016,” January 26, 2017.

Clark, Kathleen. “The Architecture of Accountability: A Case Study of the Warrantless Surveillance Program.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 12, 2010.

Clement, Andrew. “IXmaps - Tracking Your Personal Data through the NSA’s Warrantless Wiretapping Sites.” IEEE International Symposium on Technology and Society, 2013.

Coaffee, Jon, and Pete Fussey. “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, no. 1 (February 2015): 86–105.

Cohn, Cindy. “Witness Statement of Cindy Cohn in Big Brother Watch v. United Kingdom, No. CC1.” European Court of Human Rights, September 27, 2013.

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

Currie, James T. “Iran-contra and Congressional Oversight of the CIA.” *International Journal of Intelligence and CounterIntelligence* 11, no. 2 (June 1998): 185–210.

Currie, Nicholas, and Ken Stiefvater. "Counter-Terrorism Technology Assessment and Methodology Study." Final Technical Report. Rome: Air Force Research Laboratory Information Directorate, May 2003.

Davies, P.H.J. "Ideas of Intelligence: Divergent Concepts and National Institutions." *Harvard International Review* 24, no. 3 (Fall 2002): 62–66.

———. "Intelligence Culture and Intelligence Failure in Britain and the United States." *Cambridge Review of International Affairs* 17, no. 3 (October 2004): 495–520.

Degli Esposti, Sara, and Elvira Santiago-Gomez. "Acceptable Surveillance-Orientated Security Technologies: Insights from the SurPRISE Project." *Surveillance & Society* 13, no. 3/4 (2015): 437–54.

Degli Esposti, Sara, Vincenzo Pavone, and Elvira Santiago-Gomez. "Aligning Security and Privacy: The Case of Deep Packet Inspection." In *Surveillance, Privacy and Security: Citizens' Perspectives*, edited by Michael Friedewald, 71–90. PRIO New Security Studies. London; New York: Routledge, Taylor & Francis Group, 2017.

De Hert, Paul. "Balancing Security and Liberty within the European Human Rights Framework." *Utrecht Law Review*, 68-96, 1, no. 1 (September 2005).

DeLong, John, and Susan Hennessey. "Understanding Footnote 14: NSA Lawyering, Oversight, and Compliance." Lawfare, October 7, 2016. https://www.lawfareblog.com/understanding-footnote-14-nsa-lawyering-oversight-and-compliance.

Department of Justice, Office of Director of National Intelligence. "Background Paper on Title VII of FISA." Appendix to Senate Committee on Intelligence Report on FAA Sunsets Extension Act of 2012, June 7, 2012.

Derix, Steven, Glenn Greenwald, and Huib Modderkolk. "Dutch Intelligence Agency AIVD Hacks Internet Forums." *Nrc.Nl*, November 30, 2013. https://www.nrc.nl/nieuws/2013/11/30/dutch-intelligence-agency-aivd-hacks-internet-fora-a1429280.

DeSilver, Drew. "Most Young Americans Say Snowden Has Served the Public Interest." Pew Research Center, January 22, 2014. http://www.pewresearch.org/fact-tank/2014/01/22/most-young-americans-say-snowden-has-served-the-public-interest/.

Dietrich, Jan-Hendrik. "Of Toothless Windbags, Blind Guardians and Blunt Swords: The Ongoing Controversy about the Reform of Intelligence Services Oversight in Germany." *Intelligence and National Security* 31, no. 3 (April 15, 2016): 397–415.

Diffie, Whitfield, and Susan Eva Landau. *Privacy on the Line: The Politics of Wiretapping and Encryption*. Updated and expanded ed. Cambridge, Mass: MIT Press, 2007.

Dillon, Thomas W., and Daphyne S. Thomas. "Exploring the Acceptance of Body Searches, Body Scans and TSA Trust." *Journal of Transportation Security* 8, no. 3–4 (December 2015): 51–67.

Director of Legislative Affairs, Office of the Director of National Intelligence, and Assistant Attorney General. "The Intelligence Community's Collection Program Under Title VII of the Foreign Intelligence Surveillance Act," May 4, 2012.

Ditton, Jason, and Emma Short. "Yes, It Works, No, It Doesn't: Comparing the Effects of Open CCTV in Two Adjacent Scottish Town Centres." *Crime Prevention Studies* 10 (1999): 201–24.

Doherty, Carroll. "Balancing Act: National Security and Civil Liberties in Post-9/11 Era." Pew Research Center, June 7, 2013. http://www.pewresearch.org/fact-tank/2013/06/07/balancing-act-national-security-and-civil-liberties-in-post-911-era/.

Drakos, Konstantinos, and Nicholas Giannakopoulos. "An Econometric Analysis of Counterterrorism Effectiveness: The Impact on Life and Property Losses." *Public Choice* 139, no. 1–2 (April 2009): 135–51.

Edwards, Matthew, Awais Rashid, and Paul Rayson. "A Systematic Survey of Online Data Mining Technology Intended for Law Enforcement." *ACM Computing Surveys* 48, no. 1 (September 2015).

Ekblom, Paul. *Crime Prevention, Security and Community Safety Using the 5Is Framework*. Houndmills, Basingstoke; New York, NY: Palgrave Macmillan, 2011.

Eskens, Sarah, Ot van Daalen, and Nico van Eijk. "Ten Standards for Oversight and Transparency of National Intelligence Services." Institute for Information Law, University of Amsterdam, 2015.

Farrington, David, Martin Gill, Sam Waples, and Javier Argomaniz. "The Effects of Closed-Circuit Television on Crime: Meta-Analysis of an English National Quasi-Experimental Multi-Site Evaluation." *Journal of Experimental Criminology* 3 (March 1, 2007): 21–38.

Farson, Stuart, and Reg Whitaker. "Accounting for the Future or the Past?: Developing Accountability and Oversight Systems to Meet Future Intelligence Needs." In *The Oxford Handbook of National Security Intelligence*, edited by Loch K. Johnson. Oxford Handbooks. Oxford; New York: Oxford University Press, 2010.

Ford, Christopher. "Intelligence Demands in a Democratic State: Congressional Intelligence Oversight." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, July 1, 2006. https://papers.ssrn.com/abstract=2628680.

Friedewald, Michael, and Ronald J. Pohoryles. "Technology and Privacy." *Innovation: The European Journal of Social Science Research* 26, no. 1–2 (March 2013): 1–6.

Fuchs, Christian, and Daniel Trottier. "Internet Surveillance after Snowden: A Critical Empirical Study of Computer Experts' Attitudes on Commercial and State Surveillance of the Internet and Social Media Post-Edward Snowden." *Journal of Information, Communication and Ethics in Society* 15, no. 4 (November 13, 2017): 412–44.

Fuentes, Luis M. and Sergio A. Velastin. "Tracking-based event detection for CCTV systems." *Pattern Analysis and Applications* 7, no. 4 (November 26, 2004): 356-64.

Fulgham, David A. "Pressure Mounts to Find New Science to Meet Cyber-Intel Needs." *Aviation Week & Space Technology*, March 29, 2010. https://aviationweek.com/awin/pressure-mounts-find-new-science-meet-cyber-intel-needs.

Gallagher, Sean. "Building a Panopticon: The Evolution of the NSA's XKEYSCORE," *Ars Technica*, August 9, 2013. http://arstechnica.com/information-technology/2013/08/building-a-panopticon-the-evolution-of-the-nsas-xkeyscore/.

Gao, George. "What Americans Think about NSA Surveillance, National Security and Privacy." Pew Research Center, May 29, 2015. http://www.pewresearch.org/fact-tank/2015/05/29/what-americans-think-about-nsa-surveillance-national-security-and-privacy/.

Gill, Martin, Anthea Rose, Kate Collins, and Martin Hemming. "Redeployable CCTV and Drug-Related Crime: A Case of Implementation Failure." *Drugs: Education, Prevention and Policy* 13, no. 5 (January 1, 2006): 451–60.

Gill, Martin, and Angela Spriggs. "Assessing the Impact of CCTV." Home Office Research Study. Home Office Research, Development and Statistics Directorate, February 2005. https://www.cctvusergroup.com/downloads/file/Martin%20gill.pdf.

Gill, Peter. *Intelligence Governance and Democratisation: A Comparative Analysis of the Limits of Reform*. Studies in Intelligence. New York, NY: Routledge, 2016.

Graaff, Bob de, and James M. Nyce, eds. "Introduction." In *Handbook of European Intelligence Cultures*. Lanham: Rowman & Littlefield Education, A division of Rowman & Littlefield Publishers, Inc., 2016.

Graaf, Gjalt de, Leo Huberts, and Remco Smulders. "Coping with Public Value Conflicts." *Administration & Society* 48, no. 9 (November 2016): 1101–27.

Graaf, Gjalt de, and Zeger van der Wal. "Managing Conflicting Public Values: Governing with Integrity and Effectiveness." *The American Review of Public Administration* 40, no. 6 (November 2010): 623–30.

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

Greenwald, Glenn, and Ewen MacAskill. "NSA Prism Program Taps in to User Data of Apple, Google and Others." *The Guardian*, June 7, 2013. http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data.

Guelke, John, Tom Sorell, Katerina Hadjimatheou, Martin Scheinin, Jonathan Andrew, Juha Lavapuro, Tuomas Ojanen, et al. "SURVEILLE Deliverable 2.6: Matrix of Surveillance Technologies." Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme, July 31, 2013.

Guerrier, Claudine. *Security and Privacy in the Digital Era*. John Wiley & Sons, 2016.

Gunasekara, Gehan, Andrew A. Adams, and Kiyoshi Murata. “Ripples down under: New Zealand Youngsters’ Attitudes and Conduct Following Snowden.” *Journal of Information, Communication and Ethics in Society* 15, no. 3 (August 14, 2017): 297–310.

Hayden, Michael. “Balancing Security and Liberty: The Challenge of Sharing Foreign Signals Intelligence.” *Notre Dame Journal of Law, Ethics & Public Policy* 19, no. 1 (February 3, 2014): 247.

———. Munk Debates - State Surveillance - Pre-Debate Interview, 2014. https://munkdebates.com/debates/state-surveillance.

———. “Remarks at the Atlantic Council,” November 13, 2008.

———. “Remarks at the Council on Foreign Relations,” September 7, 2007.

———. “Remarks by General Michael V. Hayden: What American Intelligence & Especially the NSA Have Been Doing to Defend the Nation.” National Press Club, January 23, 2006.

———. “Statement for the Record before the Joint Inquiry of the Senate Select Committee on Intelligence and the House Permanent Select Committee on Intelligence,” 2002.

Hayden, Michael V. “BEYOND SNOWDEN: An NSA Reality Check.” *World Affairs* 176, no. 5 (2014): 13–23.

———. “To Keep America Safe, Embrace Drone Warfare.” *The New York Times*, February 19, 2016, sec. Opinion. https://www.nytimes.com/2016/02/21/opinion/sunday/drone-warfare-precise-effective-imperfect.html.

Head of the Interception of Communications Commissioner’s Office. “Circular to All Senior Responsible Officers under Chapter 2 of Part I of the Regulation of Investigatory Powers Act 2000 (RIPA 2000) Regarding Applicant Errors,” September 1, 2014.

Her Majesty’s Inspectorate of the Constabulary. “An Inspection of the National Crime Agency,” 2015. https://www.justiceinspectorates.gov.uk/hmicfrs/wp-content/uploads/an-inspection-of-the-national-crime-agency. pdf.

Hewitt, Christopher. “Law Enforcement Tactics and Their Effectiveness in Dealing with American Terrorism: Organizations, Autonomous Cells, and Lone Wolves.” *Terrorism and Political Violence* 26, no. 1 (January 2014): 58–68.

Hijzen, Constant. “More than a Ritual Dance. The Dutch Practice of Parliamentary Oversight and Control of the Intelligence Community.” *Security and Human Rights* 24, no. 3–4 (April 30, 2014): 227–38.

House Permanent Select Committee on Intelligence. Cyber Intelligence Sharing and Protection Act, Pub. L. No. H.R. 624 (2013).

———. FISA Transparency and Modernization Act, Pub. L. No. H.R. 4291 (2014).

———. Haqqani Network Terrorist Designation Act of 2012, Pub. L. No. H.R. 6036 (2012).

———. "Intelligence Authorization Act for Fiscal Year 2008 - Conference Report," December 6, 2007.

———. Intelligence Authorization Act for Fiscal Year 2014, Pub. L. No. H.R. 3381 (2013).

———. "Intelligence Authorization Act for Fiscal Year 2014 - Report," November 25, 2013.

———. Intelligence Authorization Act for Fiscal Year 2016 (2015).

———. "Intelligence Authorization Act for Fiscal Year 2016 - Report," June 9, 2015.

———. Protecting Cyber Networks Act (2015).

———. "Semi-Annual Report of the Activity of the House Permanent Select Committee on Intelligence," June 30, 2011.

Intelligence and Security Committee. "Access to Communications Data by the Intelligence and Security Agencies," February 2013.

———. "Intelligence and Security Committee Annual Report 2006-2007," January 2008.

———. "Intelligence and Security Committee Annual Report 2007-2008," March 2009.

———. "Intelligence and Security Committee Annual Report 2008-2009," March 2010.

———. "Intelligence and Security Committee Annual Report 2009-2010," March 2010.

———. "Intelligence and Security Committee Annual Report 2010-2011," July 2011.

———. "Intelligence and Security Committee Annual Report 2011-2012," July 2012.

———. "Intelligence and Security Committee Annual Report 2012-2013," July 2013.

———. "Intelligence and Security Committee Annual Report 2013-2014," November 2014.

———. "Intelligence and Security Committee Annual Report 2015-2016," July 2016.

———. "Privacy and Security: A Modern and Transparent Legal Framework," March 2015.

———. "Report on the Draft Investigatory Powers Bill," February 2016.

———. "Statement on GCHQ's Alleged Interception of Communications under the US PRISM Programme," n.d.

Intelligence Services Commissioner. “Report of the Intelligence Services Commissioner for 2006,” January 2008.

———. “Report of the Intelligence Services Commissioner for 2007,” July 2008.

———. “Report of the Intelligence Services Commissioner for 2008,” July 2009.

———. “Report of the Intelligence Services Commissioner for 2009,” July 2010.

———. “Report of the Intelligence Services Commissioner for 2010,” June 2011.

———. “Report of the Intelligence Services Commissioner for 2011,” July 2012.

———. “Report of the Intelligence Services Commissioner for 2012,” July 2013.

———. “Report of the Intelligence Services Commissioner for 2013,” June 2014.

———. “Report of the Intelligence Services Commissioner for 2014,” June 2015.

———. “Report of the Intelligence Services Commissioner for 2015,” September 2016.

Intelligence Services Commissioner, and Head of the Interception of Communications Commissioner’s Office. Joint Committee on the Draft Investigatory Powers Bill Oral evidence (Dec. 21, 2015).

Interception of Communications Commissioner. “2010 Annual Report of the Interception of Communications Commissioner,” June 2011.

———. “2011 Annual Report of the Interception of Communications Commissioner,” July 2012.

———. “2012 Annual Report of the Interception of Communications Commissioner,” July 2013.

———. “2013 Annual Report of the Interception of Communications Commissioner,” April 2014.

———. “Evidence for Investigatory Powers Review,” December 5, 2014.

———. “Evidence for the Joint Committee for the Investigatory Powers Bill,” December 21, 2015.

———. “Half-Yearly Report of the Interception of Communications Commissioner July 2015,” July 2015.

———. Oral evidence (2014).

———. “Report of the Interception of Communications Commissioner Annual Report for 2015,” September 2016.

———. “Report of the Interception of Communications Commissioner for 2005-2006,” February 2007.

———. “Report of the Interception of Communications Commissioner for 2006,” January 2008.

———. “Report of the Interception of Communications Commissioner for 2007,” July 2008.

———. “Report of the Interception of Communications Commissioner for 2008,” July 2009.

———. “Report of the Interception of Communications Commissioner for 2009,” July 2010.

———. “Report of the Interception of Communications Commissioner for 2014,” March 2015.

———. Supplementary evidence to oral session on 4th November 2014 (2014).

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, 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), 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), September 9, 2015.

Investigatory Powers Tribunal. Abdel-Hakim Belhaj & others v. Security Service & others, No. [2015] UKIPTrib 13_132-H (Investigatory Powers Tribunal, April 29, 2015).

———. Caroline Lucas & others v. Government Communications Headquarters & others, No. [2015] UKIPTrib 14_79-CH (October 14, 2015).

———. Human Rights Watch Inc & Ors v. The Secretary of State for the Foreign & Commonwealth Office & Ors, No. [2016] UKIPTrib15_165-CH (Investigatory Powers Tribunal May 16, 2016).

———. “Investigatory Powers Tribunal Report 2010,” n.d.

———. “Investigatory Powers Tribunal Report 2011-2015,” 2016.

———. Liberty v. Government Communications Headquarters & Others Amended Determination, No. [2015] UKIPTrib 13_77-H_2 (Investigatory Powers Tribunal June 22, 2015).

———. Liberty v. The Government Communications Headquarters & Others, No. [2014] UKIPTrib 13_77-H (Investigatory Powers Tribunal December 5, 2014).

———. Privacy International & others v. The Secretary of State for Foreign and Commonwealth Affairs & The Government Communications Headquarters, No. [2016] UKIPTrib 14_85-CH (Investigatory Powers Tribunal December 12, 2016).

———. Privacy International v. Secretary of State for Foreign and Commonwealth Affairs & others, No. [2016] UKIPTrib 15_110-CH (Investigatory Powers Tribunal October 17, 2016).

Johnson, Loch K. *A Season of Inquiry: The Senate Intelligence Investigation*, 2014.

Johnson, Loch K. "The U.S. Congress and the CIA: Monitoring the Dark Side of Government." *Legislative Studies Quarterly* 5 (November 1980): 477–99.

———. "Governing in the Absences of Angels: On the Practice of Intelligence Accountability in the United States." In *Who's Watching the Spies? Establishing Intelligence Service Accountability*, edited by Hans Born, Lock K. Johnson, and Ian Leigh, 1st ed., 57–78. Washington, DC: Potomac Books, 2005.

———. "The Church Committee Investigation of 1975 and the Evolution of Modern Intelligence Accountability." *Intelligence and National Security* 23, no. 2 (April 2008): 198–225.

———. "The Contemporary Presidency: Presidents, Lawmakers, and Spies: Intelligence Accountability in the United States." *Presidential Studies Quarterly* 34, no. 4 (December 2004): 828–37.

Jonas, Jeff, and Jim Harper. "Effective Counterterrorism and the Limited Role of Predictive Data Mining." Policy Analysis. Washington D.C.: Cato Institute, December 11, 2006.

Judge, Foreign Intelligence Surveillance Court. "Amendment to Order for Purposes of Querying the Metadata Archive, BR 07-10," May 31, 2007.

———. "Amendment to Primary Order, BR 11-07," February 10, 2011.

———. "Memorandum Opinion," October 3, 2011.

———. "Memorandum Opinion," November 30, 2011.

———. "Memorandum Opinion," 2012.

———. "Memorandum Opinion and Order," April 26, 2017.

———. "Order, BR 06-08," August 18, 2006.

———. "Order, BR 06-12," November 15, 2006.

———. “Order, BR 07-04,” February 7, 2007.

———. “Order, BR 07-10,” May 3, 2007.

———. “Order, BR 07-14,” July 25, 2007.

———. “Primary Order, BR 07-16,” October 18, 2007.

———. “Primary Order, BR 08-01,” January 2008.

———. “Primary Order, BR 08-04,” April 3, 2008.

———. “Primary Order, BR 08-07,” June 26, 2008.

———. “Primary Order, BR 08-08,” August 19, 2008.

———. “Primary Order, BR 08-13,” December 12, 2008.

———. “Primary Order, BR 09-01,” March 5, 2009.

———. “Primary Order, BR 09-06,” May 29, 2009

———. “Primary Order, BR 10-10,” February 26, 2010.

———. “Primary Order, BR 10-17,” May 14, 2010.

———. “Primary Order, BR 10-49,” August 4, 2010.

———. “Primary Order, BR 10-70,” October 29, 2010.

———. “Primary Order, BR 11-07,” January 20, 2011.

———. “Primary Order, BR 11-57,” April 13, 2011.

———. “Primary Order, BR 11-107,” June 22, 2011.

———. “Supplemental Opinion,” December 12, 2008.

———. “Supplemental Order, BR 11-57,” April 13, 2011.

———. “Supplemental Order, BR 11-107,” June 22, 2011.

Kavathatzopoulos, Iordanis, Ryoko Asai, Andrew A. Adams, and Kiyoshi Murata. “Snowden’s Revelations and the Attitudes of Students at Swedish Universities.” *Journal of Information, Communication and Ethics in Society* 15, no. 3 (August 14, 2017): 247–64.

Kibbe, Jennifer. “Congressional Oversight of Intelligence: Is the Solution Part of the Problem?” *Intelligence and National Security* 25, no. 1 (February 2010): 24–49.

Knott, Stephen. “Executive Power and the Control of American Intelligence.” *Intelligence and National Security* 13, no. 2 (June 1998): 171–76.

Kraan, Jeroen. “Achtergrond: Dit Moet Je Weten over de ‘Aftapwet’ En Het Referendum.” *Nu.nl*, October 9, 2017. https://www.nu.nl/internet/4956455/achtergrond-moet-weten-aftapwet-en-referendum.html.

Kreiger, W. “Oversight of Intelligence: A Comparative Approach.” In *National Intelligence Systems: Current Research and Future Prospects*, edited by Gregory F. Treverton and Wilhelm Agrell, 210–34. New York: Cambridge University Press, 2009.

Landau, Susan. “Security and Privacy: Facing Ethical Choices.” *IEEE Security & Privacy* 12, no. 4 (July 2014): 3–6..

Lawfare Podcast. *Inside NSA, Part I--An Interview with General Counsel Rajesh De*. Inside NSA. Accessed November 10, 2017. https://www.lawfareblog.com/lawfare-podcast-episode-52-inside-nsa-part-i-interview-general-counsel-rajesh-de.

———. *Inside NSA, Part II--Wherein We Interview the Agency’s Chief of Compliance, John DeLong*. Inside NSA. Accessed November 10, 2017. https://www.lawfareblog.com/lawfare-podcast-episode-53-inside-nsa-part-ii-wherein-we-interview-agencys-chief-compliance-john.

Leigh, Ian. “More Closely Watching the Spies: Three Decades of Experiences.” In *Who’s Watching the Spies? Establishing Intelligence Service Accountability*, edited by Hans Born, Lock K. Johnson, and Ian Leigh, 1st ed., 3–12. Washington, DC: Potomac Books, 2005.

———. “The Accountability of Security and Intelligence Agencies.” In *Handbook of Intelligence Studies*, edited by Loch K. Johnson, 67–81. New York/ London: Routledge, 2007.

Lingel, Sherrill Lee, Lance Menthe, Brien Alkire, John Gibson, Scott A. Grossman, Robert A. Guffey, Keith Henry, Lindsay D. Millard, and Christopher Mouton. *Methodologies for Analyzing Remotely Piloted Aircraft in Future Roles and Missions*. Documented Briefing. Santa Monica, CA: RAND, 2012.

Lobban, Sir Iain. “Sir Iain Lobban’s Valedictory Speech - as Delivered,” October 21, 2014. https://webarchive.nationalarchives.gov.uk/20170303162613/https://www.gchq.gov.uk/speech/sir-iain-lobbans-valedictory-speech-delivered.

Lobban, Sir Iain, Andrew Parker, and John Sawers. “Intelligence and Security Committee of Parliament: Uncorrected Transcript of Evidence,” November 7, 2013.

Lowenthal, Mark M. *Intelligence : From Secrets to Policy*. 5th ed. Los Angeles: SAGE/CQ Press, 2012.

Lum, Cynthia, Leslie W. Kennedy, and Alison Sherley. “Are Counter-Terrorism Strategies Effective? The Results of the Campbell Systematic Review on Counter-Terrorism Evaluation Research.” *Journal of Experimental Criminology* 2, no. 4 (November 1, 2006): 489–516.

Madden, Mary. “Public Perceptions of Privacy and Security in the Post-Snowden Era.” Pew Research Center, November 2014. http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/.

Madden, Mary, and Lee Rainie. “Americans’ Attitudes About Privacy, Security, and Surveillance.” Pew Research Center, May 20, 2015. http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/05/Privacy-and-Security-Attitudes-5.19.15_FINAL.pdf.

Mansfield-Devine, Steve. “The Privacy Dilemma.” *Network Security*, February 2015, 5–10.

Marcus, J. Scott. “Declaration of J. Scott Marcus in Support of Plaintiffs’ Motion for Preliminary Injunction,” No. C-06-0672-VRW (United States District Court for the Northern District of California, June 8, 2006).

Mayer, Jane. “Thomas Drake vs. the N.S.A.” *The New Yorker*, May 16, 2011. https://www.newyorker.com/magazine/2011/05/23/the-secret-sharer.

McCubbins, Mathew D., and Thomas Schwartz. “Congressional Oversight Overlooked: Police Patrols versus Fire Alarms.” *American Journal of Political Science* 28, no. 1 (February 1984): 165.

Mitrou, Lilian. “Communications Data Retention: A Pandora’s Box for Rights and Liberties?” In *Digital Privacy: Theory, Technologies, and Practices*, 409–33. Boca Raton: Auerbach Publications, 2008.

Mols, Anouk, and Susanne Janssen. “Not Interesting Enough to Be Followed by the NSA: An Analysis of Dutch Privacy Attitudes.” *Digital Journalism* 5, no. 3 (March 16, 2017): 277–98.

Monahan, Torin. “Built to Lie: Investigating Technologies of Deception, Surveillance, and Control.” *The Information Society* 32, no. 4 (August 7, 2016): 229–40.

Monahan, Torin, and Neal A. Palmer. “The Emerging Politics of DHS Fusion Centers.” *Security Dialogue* 40, no. 6 (December 2009): 617–36.

Morgan, Samuel A. “Security vs. Liberty: How to Measure Privacy Costs in Domestic Surveillance Programs.” Master’s thesis, Naval Postgraduate School, 2014.

Mueller, John E., and Mark G. Stewart. *Terror, Security, and Money: Balancing the Risks, Benefits, and Costs of Homeland Security*. Oxford; New York: Oxford University Press, 2011.

Murata, Kiyoshi, Andrew A. Adams, and Ana María Lara Palma. “Following Snowden: A Cross-Cultural Study on the Social Impact of Snowden’s Revelations.” *Journal of Information, Communication and Ethics in Society* 15, no. 3 (August 14, 2017): 183–96.

Myhill, Andy, and Paul Quinton. “It’s a Fair Cop? Police Legitimacy, Public Cooperation, and Crime Reduction.” National Policing Improvement Agency, September 2011.

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

Nivola, Pietro S. 2012. “Learning from James Q. Wilson.” The Brookings Institute. https://www.brookings.edu/articles/learning-from-james-q-wilson/

NSA. “Tor Stinks.” June 2012.

———. “XKEYSCORE.” February 25, 2008.

NSA Director of Civil Liberties and Privacy Office. “NSA’s Civil Liberties and Privacy Protections for Targeted SIGINT Activities under Executive Order 12333,” October 7, 2014.

NSA General Counsel. “NSA General Counsel Rajesh De Speech at Georgetown.” Lawfare, February 27, 2013. https://www.lawfareblog.com/nsa-general-counsel-rajesh-de-speech-georgetown.

NSA Inspector General. “Review of the Presidents’ Surveillance Program,” March 24, 2009.

NSA Inspector General, NSA General Counsel, and NSA Director. “Report to the Intelligence Oversight Board on NSA Activities - 4Q 2011,” April 5, 2012.

———. “Report to the Intelligence Oversight Board on NSA Activities - 4Q 2012,” March 4, 2013.

Obstfeld, Maurice, Jay Shambaugh, and Alan Taylor. “The Trilemma in History: Tradeoffs among Exchange Rates, Monetary Policies, and Capital Mobility.” Cambridge, MA: National Bureau of Economic Research, March 2004.

O’Connell, Kevin M. “Thinking About Intelligence Comparatively.” *Brown Journal of World Affairs* XI, no. 1 (Summer/ Fall 2004): 189–99.

Office of the Inspector General of the Department of Defense. “Audit of the Requirements for the TRAILBLAZER and THINTHREAD Systems,” December 15, 2004.

“Office of the Inspector General (OIG) - NSA.gov.” https://www.nsa.gov/about/oig/ [accessed November 13, 2017].

Olmstead, Kenneth. “Most Americans Think the Government Could Be Monitoring Their Phone Calls and Emails.” Pew Research Center, September 27, 2017. http://www.pewresearch.org/fact-tank/2017/09/27/most-americans-think-the-government-could-be-monitoring-their-phone-calls-and-emails/.

Olmsted, Kathryn S. *Challenging the Secret Government: The Post-Watergate Investigations of the CIA and FBI*. Chapel Hill: University of North Carolina Press, 1996.

Omand, David. “Countering International Terrorism: The Use of Strategy.” *Survival* 47, no. 4 (December 1, 2005): 107–16.

———. "Evidence for the Intelligence and Security Committee of Parliament," February 7, 2014.

———. "Joint Committee on the Draft Investigatory Powers Bill - Oral Evidence," December 21, 2015.

———. "Privacy and Security Inquiry: Public Evidence Session 8 - Uncorrected Transcript of Evidence." Intelligence and Security Committee of Parliament, October 23, 2014.

———. "The Future of Intelligence: What Are the Threats, the Challenges and the Opportunities?" In *The Future of Intelligence: Challenges in the 21st Century*, edited by Isabelle Duyvesteyn, Ben de Jong, and Joop van Reijn, 1 edition. Routledge, 2014.

———. "Understanding Digital Intelligence and the Norms That Might Govern It." Global Commission on Internet Governance. Centre for International Governance Innovation and Chatham House, March 19, 2015. https://www.cigionline.org/publications/understanding-digital-intelligence-and-norms-might-govern-it.

Omand, David, Jamie Bartlett, and Carl Miller. "A Balance between Security and Privacy Online Must Be Struck." Demos, 2012. https://www.demos.co.uk/wp-content/uploads/2017/03/intelligence-Report.pdf.

Orru, Elisa. "Review of European Level Studies on Perceptions of Surveillance." SURVEILLE, September 30, 2013.

Ott, Marvin C. "Partisanship and the Decline of Intelligence Oversight." *International Journal of Intelligence and CounterIntelligence* 16, no. 1 (January 1, 2003): 69–94.

Park, Chanmin, and Taehyung Wang. "Big Data and NSA Surveillance -- Survey of Technology and Legal Issues," 516–17. IEEE, 2013. https://doi.org/10.1109/ISM.2013.103.

Patsakis, Constantinos, Athanasios Charemis, Achilleas Papageorgiou, Dimitrios Mermigas, and Sotirios Pirounias. "The Market's Response toward Privacy and Mass Surveillance: The Snowden Aftermath." *Computers & Security* 73 (March 1, 2018): 194–206.

Pavone, Vincenzo, and Sara Degli Esposti. "Public Assessment of New Surveillance-Oriented Security Technologies: Beyond the Trade-off between Privacy and Security." *Public Understanding of Science* 21, no. 5 (July 2012): 556–72.

Pauley III, William H. American Civil Liberties Union against James R. Clapper (United States District Court Southern District of New York December 27, 2013).

Pepper, Sir David. "Testimony on Iraq," December 2010.

———. "The Business of Sigint: The Role of Modern Management in the Transformation of GCHQ." *Public Policy and Administration*, 2010. https://doi.org/10.1177/0952076709347080.

Perl, Raphael. “Combating Terrorism: The Challenge of Measuring Effectiveness.” CRS Report for Congress. Congressional Research Service, March 12, 2007.

Pew Research Center. “About Half See CIA Interrogation Methods as Justified,” December 15, 2014. http://www.people-press.org/2014/12/15/about-half-see-cia-interrogation-methods-as-justified/.

———. “Few See Adequate Limits on NSA Surveillance Program,” July 26, 2013. http://www.people-press.org/2013/07/26/few-see-adequate-limits-on-nsa-surveillance-program/.

———. “Global Opinions of U.S. Surveillance,” July 14, 2014. http://www.pewglobal.org/2014/07/14/nsa-opinion/.

———. “Government Surveillance: A Question Wording Experiment,” July 26, 2013. http://www.people-press.org/2013/07/26/government-surveillance-a-question-wording-experiment/.

———. “Majority Views NSA Phone Tracking as Acceptable Anti-Terror Tactic,” June 10, 2013. http://www.people-press.org/2013/06/10/majority-views-nsa-phone-tracking-as-acceptable-anti-terror-tactic/.

———. “Most Say Monitoring Allied Leaders’ Calls Is Unacceptable,” November 4, 2013. http://www.people-press.org/2013/11/04/most-say-monitoring-allied-leaders-calls-is-unacceptable/.

———. “Public Split over Impact of NSA Leak, but Most Want Snowden Prosecuted,” June 17, 2013. http://www.people-press.org/2013/06/17/public-split-over-impact-of-nsa-leak-but-most-want-snowden-prosecuted/.

Phillips, Coretta. “A Review of CCTV Evaluations: Crime Reduction Effects and Attitudes to Its Use.” In *Surveillance of Public Space: CCTV, Street Lighting and Crime Prevention, Crime Prevention Studies*, edited by Kate Painter and Nick Tilley, 123–56. Monsey, NY: Criminal Justice Press, 1999.

Poullet, Yves. “The Fight against Crime and/or the Protection of Privacy: A Thorny Debate!” *International Review of Law, Computers & Technology* 18, no. 2 (July 2004): 251–73.

Prados, John. *The Family Jewels: The CIA, Secrecy, and Presidential Power*. First edition. Discovering America. Austin: University of Texas Press, 2013.

Preibusch, Sören. “Privacy Behaviors after Snowden.” *Communications of the ACM* 58, no. 5 (April 23, 2015): 48–55.

“Presidential Policy Directive - Signals Intelligence Activities.” Whitehouse.gov, January 17, 2014. https://www.whitehouse.gov/the-press-office/2014/01/17/presidential-policy-direc tive-signals-intelligence-activities.

Privacy and Civil Liberties Oversight Board. “Report on the Surveillance Program Operated Pursuant to Section 702 of the Foreign Intelligence Surveillance Act,” July 2, 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,” January 23, 2014.

———. “Semi-Annual Report, September 2012 - March 2013,” June 27, 2013.

Pütz, Ole. “From Non-Places to Non-Events: The Airport Security Checkpoint.” *Journal of Contemporary Ethnography* 41 (April 2012): 154–88.

Rainie, Lee. “The State of Privacy in Post-Snowden America.” Pew Research Center, September 21, 2016. http://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/.

Ratcliffe, Jerry. *Video Surveillance of Public Places*. Washington, D.C.: U.S. Dept. of Justice, Office of Community Oriented Policing Services, 2006. http://purl.access.gpo.gov/GPO/LPS70261.

Ratcliffe, Jerry H., Travis Taniguchi, and Ralph B. Taylor. “The Crime Reduction Effects of Public CCTV Cameras: A Multi-Method Spatial Approach.” *Justice Quarterly* 26, no. 4 (December 2009): 746–70.

Reddick, Christopher G., Akemi Takeoka Chatfield, and Patricia A. Jaramillo. “Public Opinion on National Security Agency Surveillance Programs: A Multi-Method Approach.” *Government Information Quarterly* 32, no. 2 (April 1, 2015): 129–41.

Regan, Priscilla M., and Torin Monahan. “Beyond Counterterrorism: Data Sharing, Privacy, and Organizational Histories of DHS Fusion Centers.” *International Journal of E-Politics* 4, no. 3 (33 2013): 1–14.

Reid, E.C. “Congressional Intelligence Oversight Evolution in Progress 1947-2005.” Naval Postgraduate School, 2005.

Roy, Jeffrey. “Secrecy, Security and Digital Literacy in an Era of Meta-Data: Why the Canadian Westminster Model Falls Short.” *Intelligence and National Security* 31, no. 1 (January 2, 2016): 95–117.

Runciman, Brian. “A Bigger Haystack.” *ITNOW* 54, no. 2 (June 2012): 36–37.

Sanders, Carrie B., Crystal Weston, and Nicole Schott. “Police Innovations, ‘Secret Squirrels’ and Accountability: Empirically Studying Intelligence-Led Policing in Canada.” *British Journal of Criminology* 55, no. 4 (July 2015): 711–29.

Schmidt, Fabian. “Tapping the World’s Fiber Optic Cables.” *DW*, June 30, 2013. http://www.dw.de/tapping-the-worlds-fiber-optic-cables/a-16916476.

Schneier, Bruce. “Attacking Tor: How the NSA Targets Users’ Online Anonymity.” *The Guardian*, October 4, 2013. http://www.theguardian.com/world/2013/oct/04/tor-attacks-nsa-users-online-anonymity.

———. “Beyond Security Theater - Schneier on Security,” November 2009. https://www.schneier.com/blog/archives/2009/11/beyond_security.html.

———. *Carry on: Sound Advice from Schneier on Security*. Hoboken, New Jersey: John Wiley & Sons, 2014.

———. "Crypto-Gram Newsletter," October 15, 2013. https://www.schneier.com/crypto-gram-1310.html.

Schwarz, Frederick A.O. "The Church Committee and a New Era of Intelligence Oversight." *Intelligence and National Security* 22, no. 2 (April 2007): 270–97.

Senate Select Committee on Intelligence. "Attempted Terrorist Attack on Northwest Airlines Flight 253," May 24, 2010.

———. "Central Intelligence Agency's Detention and Interrogation Program," December 9, 2014.

———. "FAA Sunsets Extension Act of 2012," June 7, 2012.

———. "FISA Improvements Act of 2013," November 12, 2013.

———. "Foreign Intelligence Surveillance Act of 1978 Amendments Act of 2007," October 26, 2007.

———. "Intelligence Activities Relating to Iraq Conducted by the Policy Counterterrorism Evaluation Group and the Office of Special Plans within the Office of the Under Secretary of Defense for Policy," June 5, 2008.

———. "Intelligence Authorization Act for Fiscal Year 2006," September 29, 2005.

———. "Intelligence Authorization Act for Fiscal Year 2007," May 25, 2006.

———. "Intelligence Authorization Act for Fiscal Year 2008," May 31, 2007.

———. "Intelligence Authorization Act for Fiscal Year 2009," May 8, 2008.

———. "Intelligence Authorization Act for Fiscal Year 2010," July 22, 2009.

———. "Intelligence Authorization Act for Fiscal Year 2010," July 19, 2010.

———. "Intelligence Authorization Act for Fiscal Year 2011," April 4, 2011.

———. "Intelligence Authorization Act for Fiscal Year 2012," August 1, 2011.

———. "Intelligence Authorization Act for Fiscal Year 2014," November 13, 2013.

———. "Intelligence Authorization Act for Fiscal Year 2015," July 31, 2014.

———. "Intelligence Authorization Act for Fiscal Year 2017," June 15, 2016.

———. "Postwar Findings About Iraq's WMD Programs and Links to Terrorism and How They Compare with Prewar Assessments," September 8, 2006.

———. "Prewar Intelligence Assessments About Postwar Iraq," May 31, 2007.

———. "Report of the Select Committee on Intelligence - January 2005 to December 2006," April 26, 2007.

———. "Report of the Select Committee on Intelligence - January 2007 to January 2009," March 9, 2009.

———. "Report of the Select Committee on Intelligence - January 2009 to January 2011," March 17, 2011.

———. "Report of the Select Committee on Intelligence - January 2011 to January 2013," March 22, 2013.

———. "Report of the Select Committee on Intelligence - January 2013 to January 2015," March 31, 2015.

———. "Terrorist Attacks on U.S. Facilities in Benghazi, Libya, September 11-12, 2012," January 15, 2014.

———. "Whether Public Statements Regarding Iraq by U.S. Government Officials Were Substantiated by Intelligence Information," June 5, 2008.

Shelton, Martin, Lee Rainie, and Mary Madden. "Americans' Privacy Strategies Post-Snowden." Pew Research Center, March 15, 2015. http://www.pewinternet.org/2015/03/16/americans-privacy-strategies-post-snowden/.

Shultz, Jr., Richard H. "Covert Action and Executive-Legislative Relations: The Iran-Contra Crisis and Its Aftermath." *Harvard Journal of Law and Public Policy* 12, no. 2 (1989): 449–82.

Simonite, Tom. "Circumventing Encryption Frees NSA's Hands Online." *MIT Technology Review*, September 6, 2013.

———. "NSA Leak Leaves Crypto-Math Intact but Highlights Known Workarounds." *MIT Technology Review*, September 9, 2013.

Snider, L. Britt. *The Agency and the Hill: CIA's Relationship with Congress, 1946 - 2004*. Washington, DC: Center for the Study of Intelligence, Central Intelligence Agency, 2008.

Sproles, Noel. "Measures of Effectiveness: The Standards for Success." University of South Australia, 1999.

Soar, Daniel. "How to Get Ahead at the NSA." *London Review of Books* 35, no. 20 (October 24, 2013): 16–18.

Staff. "GCHQ Targets Engineers with Fake LinkedIn Pages." *Spiegel Online*, November 11, 2013. http://www.spiegel.de/international/world/ghcq-targets-engineers-with-fake-linkedin-pages-a-932821.html.

———. “Inside TAO: Documents Reveal Top NSA Hacking Unit.” *Spiegel Online*, December 29, 2013. http://www.spiegel.de/international/world/the-nsa-uses-powerfultoolbox-in-effort-to-spy-on-global-networks-a-940969.html.

Stahl, Bernd. “Privacy and Security as Ideology.” *IEEE Technology and Society Magazine* 26, no. 1 (2007): 35–45.

Stalla-Bourdillon, Sophie, Joshua Phillips, and Mark D. Ryan. *Privacy vs. Security*. SpringerBriefs in Cybersecurity. London: Springer London, 2014. https://doi.org/10.1007/978-1-4471-6530-9.

Steenhuisen, Bauke Meindert. “Competing Public Values: Coping Strategies in Heavily Regulated Utility Industries.” Next Generation Infrastructures Foundation, 2009.

Stewart, Mark G., and John Mueller. “Cost-Benefit Analysis of Advanced Imaging Technology Full Body Scanners for Airline Passenger Security Screening.” *Journal of Homeland Security and Emergency Management* 8, no. 1 (January 16, 2011).

Tene, Omer. “Privacy: The New Generations.” *International Data Privacy Law* 1, no. 1 (February 1, 2011): 15–27.

Timberg, Craig and Ellen Nakashima. “Agreements with Private Companies Protect U.S. Access to Cables’ Data for Surveillance.” *The Washington Post*, July 7, 2013.

Trottier, Daniel. *Social Media as Surveillance: Rethinking Visibility in a Converging World*. Farnham, Surrey, England; Burlington, VT: Ashgate, 2012.

Trottier, Daniel, and David Lyon. “Key Features of Social Media Surveillance.” In *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*, 89–105. New York: Routledge, 2012.

Tsvetovat, Maksim, and Kathleen M. Carley. “On Effectiveness of Wiretap Programs in Mapping Social Networks.” *Computational and Mathematical Organization Theory* 13, no. 1 (November 8, 2006): 63–87.

Van Buren, Jelle. “From Oversight to Undersight: The Internationalization of Intelligence.” *Security and Human Rights* 24 (2013): 239.

Van den Broek, Tijs, Merel Ooms, Michael Friedewald, Marc Van Lieshout, and Sven Rung. “Privacy and Security: Citizens’ Desires for an Equal Footing.” In *Surveillance, Privacy and Security: Citizens’ Perspectives*, 1st ed., 15–35. Abingdon, Oxon; New York, NY: Routledge, 2017.

Van Dongen, Teun. “Break It Down: An Alternative Approach to Measuring Effectiveness in Counterterrorism.” Economics of Security Working Paper Series. DIW Berlin, German Institute for Economic Research, 2009. https://econpapers.repec.org/paper/diwdiweos/diweos23.htm.

———. “The Science of Fighting Terrorism: The Relation between Terrorist Actor Type and Counterterrorism Effectiveness.” PhD Thesis, Leiden University, 2015. https://openaccess.leidenuniv.nl/bitstream/handle/1887/29742/Dissertatie_Van_Dongen_omslag.pdf?sequence=3.

Van Gulijk, Coen, Simone Sillem, and Michelle Cayford. "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 Limitations, and Efficiency, September 30, 2013.

Van Lieshout, Marc, Michael Friedewald, David Wright, and Serge Gutwirth. "Reconciling Privacy and Security." *Innovation: The European Journal of Social Science Research* 26, no. 1–2 (March 2013): 119–32.

Van Riezen, Bram, and Karlijn Roex. "Counter-Terrorism in the Netherlands and the United Kingdom: A Comparative Literature Review Study." *Social Cosmos* 3, no. 1 (2012): 97–110.

Van Um, Eric and Daniela Pisoiu. "Effective Counterterrorism: What Have We Learned so Far?" Economics of Security Working Paper Series. DIW Berlin, German Institute for Economic Research, 2011. https://econpapers.repec.org/paper/diwdiweos/diweos55.htm.

Venetianer, Péter L. and Hongli Deng. "Performance evaluation of an intelligence surveillance system – A case study." *Computer Vision and Image Understanding* 114, no. 11 (November 2010): 1292-1302.

Verfaillie, Kristof, and Evelien van den Herrewegen. "Public Assessments of the Security/Privacy Trade-off: A Criminological Conceptualization." PRISMS, February 1, 2012.

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

Wegge, Njord. "Intelligence Oversight and the Security of the State." *International Journal of Intelligence and CounterIntelligence* 30, no. 4 (October 2, 2017): 687–700.

Welsh, Brandon. C., and David. P. Farrington. "Effects of Closed-Circuit Television on Crime." *The ANNALS of the American Academy of Political and Social Science* 587, no. 1 (May 1, 2003): 110–35.

Welsh, Brandon C., and David P. Farrington. "Public Area CCTV and Crime Prevention: An Updated Systematic Review and Meta-Analysis." *Justice Quarterly* 26, no. 4 (December 2009): 716–45.

Whiting, Alex. "President Reagan and Violations of the Boland Amendment." *First Principles: National Security and Civil Liberties* 12, no. 4 (1987): 1–10.

Willis, Henry H., Joel B. Predd, Paul K. Davis, and Wayne Brown. "Measuring the Effectiveness of Border Security Between Ports-of-Entry:" Product Page. RAND, 2010. https://www.rand.org/pubs/technical_reports/TR837.html.

Wills, Aidan. *Guidebook: Understanding Intelligence Oversight*. Geneva: Geneva Centre for the Democratic Control of Armed Forces (DCAF), 2007.

Wills, Aidan, and Mathias Vermeulen. “Parliamentary Oversight of Security and Intelligence Agencies in the European Union.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, June 1, 2011.

Wilson, James Q. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books, 1989.

Završnik, Aleš, and Pia Levičnik. “The Public Perception of Cyber-Surveillance Before and After Edward Snowden’s Surveillance Revelations.” *Masaryk University Journal of Law and Technology* 9, no. 2 (September 30, 2015).

Zegart, Amy, and Julie Quinn. “Congressional Intelligence Oversight: The Electoral Disconnection.” *Intelligence and National Security* 25, no. 6 (December 1, 2010): 744–66.

Acknowledgements

I would like to thank everyone who has made this PhD process and completion possible. Firstly, thank you to my professors: to Coen van Gulijk who granted me the privilege of being selecting as a PhD candidate and started me on this stimulating and fascinating course; to Pieter van Gelder whose positivity has encouraged me throughout; to Wolter Pieters, without whose timely entrance and regular, constructive input this dissertation would have seen a much longer and rockier road; and to Genserik Reniers. Thank you also to Michel van Eeten, whose brainstorming and mentorship supplied me with key support during a particularly difficult phase of the PhD. I also particularly enjoyed my bi-weekly visits with my colleague Zahra Rezvani, who shared the challenges of mothering while endeavoring to carry out PhD studies.

Thank you to Tamara van Dijk and her family, who provided key support via babysitting during the early years of my studies. Thank you to my dear friend Amber LaBounty for her life support. A very special thanks to Caleb Bringhurst who, at the age of 12, designed the cover for this printed book, and to his parents for their creative brainstorming and timely and generous outpouring to our family in the midst of the COVID-19 pandemic and the very final stage of my dissertation preparation. I would like to thank to my parents without whose investment in all the preceding years of my education and never-failing support and belief in me, I would have never made it to this point. And while I cannot exactly thank our two beautiful children for their contribution to this dissertation, I would like to acknowledge the richness and depth that their entrance into the world during the course of my PhD has leant to this work. Lastly, but most importantly, I wish to thank my wonderful husband, whose networking skills introduced to me the possibility of this doctorate, and whose constant support and love has seen me through to the end.

Curriculum Vitae

Michelle Cayford
Born in Spokane WA, USA on April 10, MCMLXXVII

EDUCATION

Master of Arts in International Affairs, specialization International Security
Sciences Po | Paris, France
2008 – 2010

Bachelor of Arts, History & French (*magna cum laude*)
University of Washington | Seattle, WA
1995 – 1999

PROFESSIONAL EXPERIENCE

Policy Planning Unit, Office of Secretary General (Internship) 2012
NATO | Brussels, Belgium

Researched and analyzed core issues pertaining to the direction of NATO and co-authored policy papers as both lead author and supporting author, resulting in new policy proposals and endorsement by the Secretary General. Synthesized articles and committee meetings for the Unit and Secretary General. Briefed Unit on NATO member countries' positions. Mapped U.S. Congress positions on salient issues, including NATO enlargement, burden sharing, and flexible alliance; conducted statistical analysis and charting of Members' defense budget contributions to NATO.

Consultant 2010-2011
Boislandry Consulting | Paris, France

Researched and evaluated security situations and risk potential for private clients with international locations. Monitored news and wrote press reports for subjects of importance related to clients' interests, such as regional political stability. Translated and edited intelligence products and communications from French to English.

Lead Criminal Intelligence Specialist 2005-2008
Northwest HIDTA | Seattle, WA

Provided intelligence support for federal, state, and local agencies investigating drug-trafficking organizations. Coordinated investigations through identifying duplicated efforts and liaising agencies. Co-led coordination and analysis of intelligence on regional project targeting transnational criminal organizations. Leveraged all-source data and employed techniques such as link analyses leading to dismantling of criminal networks and location and prosecutions of suspects. Drafted strategic research for threat assessments. Commended for written reports synthesizing and summarizing significant analytical findings.

Publications

Cayford, Michelle. 2015. "Measures of Success: Developing a Method for Evaluating the Effectiveness of Surveillance Technology." European Intelligence and Security Informatics Conference. 187-187. DOI:10.1109/EISIC.2015.33.

Cayford, Michelle. 2014. "SURVEILLE Paper on Mass Surveillance by the National Security Agency of the United States of America." Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme.

Cayford, Michelle, Coen van Gulijk, Erik Kempel, Juha Lavapuro, Tuomas Ojanen, Martin Scheinin, John Guelke, et al. 2015. "SURVEILLE Deliverable D2.9: Consolidated Survey of Surveillance Technologies." Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme.

Cayford, Michelle, Coen van Gulijk and Simone Sillem. 2014. "SURVEILLE Deliverable 3.8: Report Combining Results of all Effectiveness Research." Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme.

Cayford, Michelle, Coen van Gulijk, and P.H.A.J.M. van Gelder. 2014. "All swept up: An initial classification of NSA surveillance technology." In T. Nowakowski, M. Mlynczak, A. Jodeiko-Pietruczuk & S. Werbinska-Wojchiechowska (eds.) *Safety and Reliability: Methodology and Applications.* CRC Press/Balkema. 643-650. European Safety and Reliability Conference, Wroclaw, Poland. DOI: 10.1201/b17399-90.

Cayford, Michelle, Coen van Gulijk, and P.H.A.J.M. Gelder. 2014. "When Counting is Not Enough: Limitations of NSA's Effectiveness Assessment of Surveillance Technology." IEEE Joint Intelligence and Security Informatics Conference. 333-333. DOI: 10.1109/JISIC.2014.80.

Cayford, Michelle, Simone Sillem, Pei-Hui Lin, and Bert Kooij. 2015. "SURVEILLE Deliverable 3.9: Final Report of WP3." Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme.

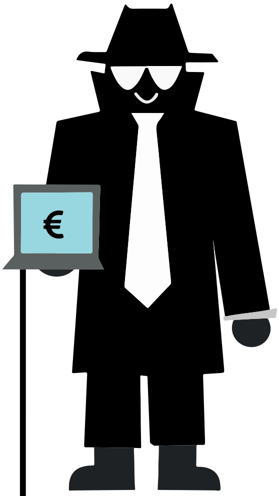
Cayford, Michelle and Wolter Pieters. 2018. "The Effectiveness of Surveillance Technology: What intelligence officials are saying." *The Information Society*, 34:2, 88-103. DOI:10.1080/01972243.2017.1414721.

Cayford, Michelle, Wolter Pieters, and Constant Hijzen. 2018. "Plots, Murders, and Money: Oversight bodies evaluating the effectiveness of surveillance technology." *Intelligence and National Security*, 33:7, 999-1021. DOI:10.1080/02684527.2018.1487159.

Cayford, Michelle, Wolter Pieters, and P.H.A.J.M. van Gelder. 2019. “Wanting it all – public perceptions of the effectiveness, cost, and privacy of surveillance technology.” *Journal of Information, Communication, and Ethics in Society.* DOI:10.1108/JICES-11-2018-0087.

Van Gulijk, Coen, Michelle Cayford, Bert-Jan Kooij, Martin Scheinin, Mathias Vermeulen, Juha Lavapuro, Tuomas Ojanen, et al. 2014. “SURVEILLE Paper Assessing Surveillance in the Context of Preventing a Terrorist Act.” Surveillance: Ethical Issues, Legal Limitations, and Efficiency. EU Seventh Framework Programme.

Van Gulijk, Coen, Simone Sillem, and Michelle Cayford. 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 Limitations, and Efficiency. EU Seventh Framework Programme.



Surveillance of communications data is a contentious topic, typically centering on privacy vs. security questions. Central to this debate, but often overlooked, is the question of the effectiveness of the surveillance technology. This dissertation focuses on intelligence agencies in the U.S. and the U.K. and the evaluation of the effectiveness of the surveillance technology they employ. It examines three stakeholders – intelligence practitioners, oversight bodies, and the public – and how they treat the question of effectiveness, including considerations of cost and proportionality. The final study considers the role of bureaucracy and its impact on effectiveness evaluation. The dissertation concludes with reflections on additional actors in the effectiveness debate and a discussion on the use of frameworks and the issue of trust.