

How to evaluate the governance of transboundary problems? Assessing a national counterterrorism strategy

Evaluation

2017, Vol. 23(4) 389–406

© The Author(s) 2017

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1356389017733340

journals.sagepub.com/home/evi



Mirko Noordegraaf, Scott Douglas, Aline Bos and Wouter Klem

Utrecht University, The Netherlands

Abstract

Wicked problems present major challenges for evaluation as they cross the boundaries between countries, policy domains, organizations, and scientific disciplines. They force evaluators to study large networks in order to trace extended chains of cause and effect. However, beyond this category of ‘regular’ wicked problems that *cross boundaries*, there is another class of even more complex problems that truly *transcend boundaries*. Problems such as terrorism, the global financial crisis, and climate change transcend distinctions between cause and effect, local and global problems, fact and impressions. These issues represent a class of transboundary wicked problems that are not about local challenges and uncertainty, but about globally connected events and ambiguity. We draw on the literature on wickedness, ambiguity, and transboundary issues, to formulate a transboundary evaluation approach capable of assessing the governance of these extremely complex problems. The usefulness of this perspective is demonstrated through our evaluation of the Dutch national counterterrorism strategy 2011–2015. We detail: (a) the evaluation framework which provided focus in a diffuse transboundary field; (b) the evaluation principles which guided us as evaluators; and (c) the evaluation conditions which allowed us to engage and empower the actors being evaluated. These components can together bolster the relevance and competence of evaluators in a complex world.

Keywords

Counterterrorism, evaluation, principles, transboundary wickedness, wicked issues

Corresponding author:

Professor Mirko Noordegraaf, Utrecht School of Governance, Utrecht University, The Netherlands.

Email: m.noordegraaf@uu.nl

Introduction

Governments around the world face extremely difficult challenges. Citizens expect them to respond to the threats of terrorism, global financial crises and climate change. Moreover, citizens expect governments to be accountable in their efforts and open to democratic scrutiny and independent evaluation. However, the evaluation of the governance of complex problems is correspondingly difficult. In 2016, the Dutch government asked us to evaluate the national counterterrorism strategy, a policy that was targeting radicalization, extremism, and sectarian violence spanning multiple public and private actors, different geographical layers, and executed between 2011 and 2015 against the backdrop of an emerging terrorist threat. We evaluated the strategy with a multidisciplinary team of academic experts, in close cooperation with the multiple parties involved. Given our experiences in this particular evaluation, we here deepen the understanding of so called ‘wicked’ issues and develop a practical evaluation perspective capable of assessing the governance of extreme complex problems.

We begin by differentiating between regular cross-boundary wicked issues and extremely complex transboundary wicked issues. Unemployment or crime are both wicked issues, in the sense that they contain multiple actors, diverging interests, cognitive and normative complexity, and uncertainty (e.g. Head, 2008). Situations are messy, problems are hard to define, solutions are unclear, and there are value differences (e.g. Head and Alford, 2015). Terrorism is wicked in a different way – it is not only intractable in terms of contested knowledge and expertise, contested values and unpredictability, but also in terms of religious and ideological dimensions, and territorial and geopolitical struggles. In addition, it is interwoven with other wicked issues (such as refugees and climate change), and is characterized by disruptive potential, political and public attention, performative effects of communication, and global organizational and financial infrastructures. We argue that this should be understood as *transboundary* wickedness, which differs from *cross-boundary* wickedness.

We also argue that transboundary wickedness requires a different evaluation perspective. In the case of cross-boundary wickedness, multiple boundaries – between actors, practices, disciplines, etc. – have to be crossed when management approaches are developed. This can be done by initiating collaborative governance, trust and learning (e.g. Head and Alford, 2015). In the case of transboundary wickedness, however, boundaries are unclear, shifting and constantly redrawn. Collaboration, trust and learning fall short, and the realm of distrust, coalitions and imagination will be activated. This has implications for the management as well as the evaluation of transboundary issues.

In case of cross-boundary wicked issues, actors want to evaluate the phenomenon to determine (1) whether strategies and management are effective and efficient, (2) whether measures work, and (3) which new courses of action might be more effective and/or legitimate. In case of transboundary wicked issues, the phenomenon cannot strictly be evaluated, as the phenomenon, its actors and arenas, potential measures and implications are not only highly ambiguous and essentially contested, but also full of hidden processes, temporary settlements, and distrust. However, people and publics still want answers to important questions. Are counterterrorism measures legitimate? Are the climate change measures effective? Did we learn from the mistakes of the global financial crisis? More than in the case of regular wickedness, strict answers cannot be given, as effects are difficult to trace (something must be prevented), as the problem is fluid, as stakeholders are unknown and there is a pervasive lack of trust. This is important for evaluators, as they want to present sound evaluation, maintain academic authority, and reduce

the chances of being dragged into political struggles. We therefore ask ourselves: *What is an appropriate evaluation approach for the governance of transboundary problems?*

We draw on our own experiences with the evaluation of the Dutch counterterrorism strategy to identify suitable evaluation frameworks, principles and conditions. Our own evaluation is used as a *reflexive case study*: not to present findings but to discuss the challenges we faced and the lessons we learned as evaluators. First, we explore the nature of cross-boundary versus transboundary wickedness, arguing that terrorism is a transboundary problem. Next we describe the practical difficulties of evaluating counterterrorism because of this extremely complex nature. We then present our experiences in conducting the evaluation, presenting the frameworks, principles and conditions which we applied. Finally, we reflect on the evaluation perspective which emerged and how this could be relevant for the evaluation of other transboundary problems.

Defining transboundary wickedness

Regular wickedness: Cross-boundary issues

Scholars of public administration have recently renewed their interest in ‘wicked issues’ (on the basis of Rittel and Webber, 1973), formulating various helpful definitions for wicked problems (for useful overviews, e.g. Head and Alford, 2015). According to Weber and Khademian (2008), wicked problems are ‘unstructured’, ‘comprise multiple, overlapping, interconnected subsets of problems that cut across multiple policy domains and levels of government’, and are ‘relentless’. According to Head and Alford (2015):

Wicked problems are generally seen as associated with social pluralism (multiple interests and values of stakeholders), institutional complexity (the context of inter-organizational cooperation and multilevel governance), and scientific uncertainty (fragmentation and gaps in reliable knowledge, p. 716).

Although the scholarship on wickedness can seem rather abstract, its resurgence is understandable from a more practical perspective. First, it enables scholars to revitalize age-old insights into complexity, related to notions such as multiple actors, interests, values, mutual dependencies, networks, and uncertainty (Van Bueren et al., 2003; Koppenjan and Klijn, 2004; Head, 2008). Together, these elements constitute ‘wickedness’.

Second, it enables scholars to tie these insights to contemporary, urgent societal matters, such as unemployment (Baum and Mitchell 2010), health care (e.g. Ferlie et al., 2011), mental health (Hannigan and Coffey, 2011), education (Bore and Wright, 2009), environment, energy and health (Turnpenny et al., 2009). In these issues, widespread organizational and societal concerns are brought together.

Third, it enables scholars to study the implications of wickedness and possible instruments that may be helpful to public officials and public managers in responding to it, because an understanding of the elements of wickedness might offer starting points for governing and managing wickedness. This in turn enables us to renew governance and management theories. Head and Alford (2015), for example, relate wicked issues to strong governance and management implications, including frame reflection and reframing, collaboration, communication, trust, and adaptive leadership. They write:

Tackling key challenges through nonstandard processes of adaptive management and networked governance becomes more important as problems exhibit higher levels of uncertainty and stakeholder contestation, for example, where key actors take divergent approaches to problem definitions and possible solutions, p. 717.

Weber and Khademian (2008) also stress the importance of collaboration. They especially focus on managers as ‘collaborative capacity builders’, aimed at ‘building long-term collaborative problem-solving capacity’ (p. 334). Ferlie et al. (2011) discuss governance implications in health care, including cross-organizational ICTs, inter-organizational learning and lateral leadership. Williams and Arrigo (2002) discuss the ‘skills, competencies and behavior of boundary spanners’ (p. 103) aimed at managing inter-organizational relationships and ‘building inter-organizational capacity’ in the face of wicked issues that call for ‘collaboration, partnership and networking’ (p. 105).

In essence, these existing notions of wickedness boil down to problems crossing multiple value systems, involving multiple organizations and connecting multiple scientific disciplines. This type of wickedness calls for cross-boundary governance, aimed at bringing actors and stakeholders together, strengthening cooperation between them, designing processes, combining values and/or enhancing horizontal learning. In response, an appropriate cross-boundary evaluation perspective traces the governance of problems across organizational limits, national borders, and disciplinary demarcations (Lawrence, 2011; Blackman et al., 2013; Walton, 2016).

Extreme wickedness: Transboundary issues

However, there are more extreme wicked problems which go further than crossing the boundaries between agencies, interests and domains. Levin et al. (2012), for example, have analyzed the management – or what they call ‘tragic governance’ – of climate change. They see climate change as a ‘super wicked’ problem as it comprises ‘four key features: time is running out; those who cause the problem also seek to provide a solution; the central authority needed to address it, is weak or non-existent; and, partly as a result, policy responses discount the future irrationally’. Although we understand what these authors mean and stress the importance of ‘super wickedness’ as well, we feel these features are too eclectic. We need a more systematic perspective on the dimensions of ‘super wickedness’, setting it apart from regular ‘wickedness’. In case of the features Levin et al. (2012) mention, it is not always clear why a feature belongs to ‘super wickedness’ and why not, and how it differs from normal ‘wickedness’.

In our view, transboundary wicked issues are distinct. Not so much because there are diverging interests and uncertainties – such as ‘substantive, strategic and institutional uncertainties’ (cf. Koppenjan and Klijn, 2013) – but because there are intense ambiguities and as a result, boundaries are confounded. Transboundary wicked issues go beyond borders and boundaries, as they are situated in transnational spaces, with overlapping jurisdictions, inter-linked policy spheres, fuzzy or ambivalent ethical standards, and much contestation, threat, aggression, emotion and fear, as well as time pressure. Boundaries are unclear, shifting and perpetually redrawn.

These wicked problems collapse distinctions between local and global, cause and consequence, past and future, liberty and repression. We call this *transboundary* wickedness, related to transboundary theory which mainly originated in the literature on risks and crisis (e.g. Boin, 2009; Ansell et al., 2010). Analytically, this is not so much rooted in notions such as networks

and uncertainty, but notions such as ambiguity and complexity (e.g. on the basis of March and Olsen, 1976; Christensen and Lægreid, 2012), surprise and imagination (Weick, 2005; Harris et al., 2010) and chaos (Williams and Arrigo, 2002; Haynes, 2007; Bore and Wright, 2009). Set against this background, we highlight three crucial elements of transboundary wickedness, when we apply wickedness theory to issues such as (counter)terrorism. As we will discuss in the next section, these three elements have a profound impact for the role and relevance of evaluations.

Transcending boundaries between action and reaction. Transboundary issues confound the separation between cause and effect. For example, the impact of the actions of terrorists and the reactions of governments cannot be disentangled. The impact of a single attack is actually felt by citizens through the repressive and expansive response of the government security apparatus, introducing spot checks, tightening airport security, etc. (Lum et al., 2006; Crenshaw, 2010; De Graaf, 2010; Schmid, 2013). The problems of terrorism and counterterrorism cannot be captured in classic sequences of action and reaction, cause and consequence. Boundaries between policy domains, global and local, action and reaction are not just merely crossed, but truly transcended as action-reaction becomes meshed into constant reaction. They transcend the distinctions between logical categories, forcing observers to replace evaluation concepts based on neat distinctions between cause and effect or input and output by transboundary concepts which can capture the dynamic of the field. The collapse of cause and effect means that we have to combine insights coming from complexity theory (Walton, 2016), as well as transnational action, networks and governance (Djelic and Sahlin-Andersson, 2006) and transboundary issues and crises (Boin, 2009; Boin et al., 2014).

Transcending boundaries between spaces. Transboundary problems also transcend the separation between geographical spaces. For example, although the World Trade Center was identified as the Ground Zero of the War on Terror, many other locations were actually simultaneously epicenters of tension: Afghanistan where the Taliban hosted Al-Qaida, Jerusalem as the rallying point of religious extremists, the American military bases in Saudi-Arabia as the topic of contention, Hamburg as the safe house of the attackers, American gas stations as the cash machine for oil-based regimes in the Middle East, etc. Unlike a pandemic, which can be traced back to a Patient Zero, terrorism is a transboundary wicked problem which occurs simultaneously and instantaneously around the world.

This means that all organizational partners are crucial to the solution, rather than being neatly divided as central and peripheral partners. For evaluation this means that there is no focal point or clearly demarked space to evaluate, only a global maelstrom of events that can only be related with some kind of precision in hindsight.

Transcending boundaries between knowledge systems. Regular wicked issues are complex because they contain cognitive and normative multiplicity and unpredictability. Several disciplines and different belief systems will have different opinions on classical wicked issues. Yet issues like terrorism are not only difficult to tackle in terms of contested knowledge and expertise. They are also intractable because they relentlessly intertwine different disciplines, logics and knowledge systems. For example, the Islamic State in Iraq and Syria (ISIS) advocates an integrated package of religious, political, economic, cultural, and epistemological perspectives on the world, just as the liberal democratic reply to terrorism intertwines ideological, scientific, and philosophical positions.

Box 1. Two types of wickedness.

	Type I Cross-boundary wickedness	Type II Transboundary wickedness
Examples	<ul style="list-style-type: none">• Unemployment• Labor market policies• Immigration measures	<ul style="list-style-type: none">• Terrorism• Climate change• Global financial crisis
Challenges	<ul style="list-style-type: none">• Multiple values• Multiple organizations• Multiple views on cause and effect	<ul style="list-style-type: none">• Entanglement of beliefs, facts, and values• Entanglement of core and periphery organizations• Entanglement of cause and effect
Responses	<ul style="list-style-type: none">• Bridging values and brokering deals• Connecting organizations to create coordination• Intertwining scientific insights to study action and reaction	<ul style="list-style-type: none">• Robust conceptualization to structure value debates• Create joint spaces for organizations to nurture contestation• Multidisciplinary theorization to study entangled interaction from all angles

This makes the ideas of terrorists so terrifying; they constitute an existentialist challenge to everything we know, believe, and desire. Characteristically for transboundary issues, however, these knowledge systems do not live in isolation of each other but actively feed and respond to each other. The different arenas are fundamentally connected. We see transboundary wicked issues not as post-boundary wickedness; borders are not eliminated but constantly reformed, reinforced, and redrawn. As a consequence, the appropriate response to transboundary wickedness should not only include collaboration and learning prescribed for cross-boundary issues, but also involve surprise, resilience, and imagination (Weick, 2005; Termeer et al., 2013).

In Box 1 we contrast cross-boundary wickedness and transboundary wickedness. We map both the challenges and required responses for each of these type of issues. On the whole, we define transboundary wicked issues as:

Issues that manifest themselves simultaneously in sectors, domains and geographies, belying distinctions between core and peripheral stakeholders, and transcending distinctions between categories of knowledge by meshing logics, value systems, and beliefs.

The challenges of evaluating a transboundary issue

If terrorism is an archetypal example of a transboundary wicked problem, poignantly representing the characteristics we discussed above, this generates several problems for the evaluation of the government response. For example, the government in the Netherlands drafted a comprehensive five-year strategy to deal with radicalization, extremism, and terrorist violence in 2011. This strategy incorporates actions from public and private actors across levels of government, and contains instruments whose effectiveness is still hotly debated among scientists. We were nevertheless asked in 2015 by the central government to assess the impact of this strategy. We describe the difficulties we encountered here.

Dutch Counterterrorism Strategy 2011–2015

The Dutch Counterterrorism Strategy 2011–2015 was chiefly authored by the Dutch Coordinator for Counterterrorism, a coordinating agency which works closely with all relevant actors and organizations in the field. The strategy itself is a 140 page document containing three key components: First, the strategy positioned the phenomenon of terrorism as a complex problem. The strategy emphasized that multiple types of terrorism can threaten liberal democracies, that people can be radicalized through different circumstances, and that a heavy-handed government response can spark more fear and so cause the greatest damage to society. Second, the strategy advocated a comprehensive approach, incorporating preventive and repressive interventions against radicalization and violent extremism. Third, the strategy stressed the importance of integral cooperation between all the different public and private organizations involved, spanning from the police and social services to faith groups and telecom companies. See Figure 1 for an overview of the main assumptions in the strategy.

At the confirmation of the strategy in 2011, the Dutch parliament requested that the policy would be evaluated in 2015. The Scientific Research and Documentation Center of the Ministry of Justice commissioned Utrecht University to conduct the evaluation. We assembled a team of scientific experts with practical evaluation experience; four researchers with expertise in governance, public management, and terrorism, supported by another layer of experts in constitutional law, radicalization, crisis management, and collaboration. The evaluation was also supported by an advisory board, consisting of several academics and representatives of the representatives of the National Coordinator for Counterterrorism and the Public Prosecution. As a team, we were confronted with several complications in the evaluation, stemming from the transboundary nature of the subject.

Contested expertise

We were faced with an extremely ambiguous expert base, which means it is very hard if not impossible to find common ground for methods, criteria, or judgements. For a start, it is unclear who ranks as ‘expert’. Because terrorism is shaped by psychological, social, technological, geopolitical, and financial processes, there are many people who claim to be knowledgeable about terrorism. Yet when it comes to providing authoritative insight and/or judgement, their authority will be heavily contested by other disciplines. These various experts have different disciplinary backgrounds and different disciplines understand the issue(s) differently. Evaluating counterterrorism from psychological, social or technological perspectives sheds radically different light upon the phenomenon and gives rise to highly different assessments. To top it off, expertise might be interwoven with organizational interests, especially when experts work for or are financed by government agencies. It is crucial to activate all available expertise for the evaluation, but this means that ‘independent’ assessments are difficult to find.

Conflicted ideologies

There are fundamental values at stake. Not only values relating to evaluation quality, such as truthfulness or dignity, but fundamental values, such as freedom, humanity, and the rule of law. Because transboundary wicked issues have many threatening elements, it is attractive to

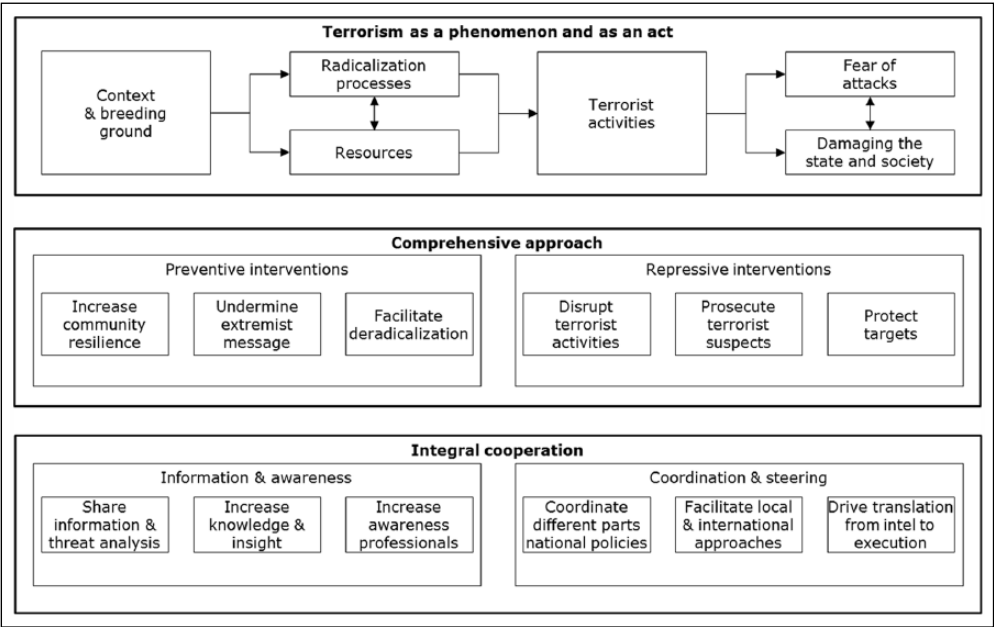


Figure 1. Policy assumptions within the counterterrorism strategy 2011–2015.
Source: Noordegraaf et al. (2016).

tackle them by forming strong policies and taking radical measures. Especially when attacks are near or when they have taken place, actors tend to ask for strong and ‘violent’ measures. This creates many paradoxes, such as, the reduction of freedoms by government to prevent the attack of freedom by terrorists. These political conflicts spill over into the evaluation, made more intense by the fact that effectiveness is hard to prove.

Moreover, these issues are discussed in international, transnational and supranational forums, and they will be influenced by other (geopolitical) debates that evolve in these very same forums. In that sense, evaluation – which is never neutral and always subjected to political forces – will be captured by political forces in one way or the other. Reports and insights will be taken out of their contexts and run the risk of being exposed in ways that serve certain agendas and interests, at the expense of others. This might affect reports and insights in multiple ways, varying from neutralization to manipulation. This can make or break evaluation reputations.

Hidden processes

Many aspects of counterterrorism are difficult to trace for evaluators, both the terrorist threat and the government response remains covert. This means that certain facts and figures and specific evidence are difficult to collect, and/or that certain facts and insights are (highly) classified. One part of the hidden nature of transboundary wicked issues is caused by the phenomenon itself. The planning and execution of terrorist attacks, either by ‘lone wolfs’ or cells, is conducted through intangible processes of ‘radicalization’ both literally and figuratively beyond the sight of governmental agencies. What intelligence is collected will remain hidden

for most evaluators. This might even concern very ‘simple’ facts such as the number of people working on counterterrorism in various agencies.

Indeterminate yardsticks

It is difficult if not impossible to determine when counterterrorism governance and management are ‘successful’. First, governments try to prevent something – terrorist attacks. We might say strategies are unsuccessful when a terrorist attack takes place, but this does not imply that they are successful when there are no attacks. Moreover, it is questionable whether attacks can be linked to strategies and whether an attack says something about the presence or absence of successful governance. Even if strategies are effective, there might be an attack, just like a ‘good’ or effective medical doctor might experience the loss of one or more patients.

Second, it is questionable whether attacks are a yardstick for success. Terrorist ambitions go beyond mere attacks, they are geared towards de-stabilizing societies, this ‘theatre of fear’ is a grand process caused by intricate cause-effect links, mediated by political and public communication (De Graaf, 2010). Such communication processes are hard to evaluate objectively, and the evaluation studies themselves become part of the terrorism discourse. Even if evaluation is aimed at ‘understanding’ acts and effects, it might be drawn into political communication games and affect perceptions of success. This is the classic Thomas theorem (‘if men define situations as real, they are real in their consequences’), but in a more intense form. There is more ‘reality’ and more potential ‘consequence’.

Developing a transboundary evaluation approach

Given these challenges, we concluded that the governance of transboundary issues cannot be assessed through strict and rigid evaluation procedures. Rigid governance of extremely wicked issues is impossible, and correspondingly, rigid evaluation is impossible as well. The governance of extremely complex issues can still be evaluated, however, and evaluation can be meaningful for the multiple actors involved (Walton, 2016). Throughout our evaluation of the Dutch counterterrorism strategy we tacitly developed evaluation frameworks, principles and conditions which together constitute a practical evaluation approach for transboundary issues. This approach consisted of an evaluation framework providing a clear focus for both the evaluators and the actors to be evaluated, supported by evaluation principles guiding the outlook of the evaluators, and evaluation conditions engaging and channeling the participation of the actors being evaluated. We present these frameworks, principles, and conditions consecutively in the paragraphs below, although these components were developed and used alongside each other in practice (see Box 2).

Evaluation framework to provide shared focus

When assessing government action in the context of clear targets, such as polio vaccinations, or clearly prescribed processes, such as appeal handling at the tax office, an evaluation can be focused on the fulfillment of these targets or the compliance with procedures. When assessing government action within the context of cross-boundary issues, the focus can be on the collaborative ability of actors involved. The question is how transboundary issues can be evaluated. We cannot just look at the observance of protocols and procedures, as terrorism is about

Box 2. Components of a transboundary evaluation approach.

Evaluation framework to provide shared focus	Evaluation principles to guide evaluators	Evaluation conditions to engage actors evaluated
1. Recognize the radical complexity of contexts	1. Robust conceptualization of object to be evaluated	1. Joint space where different actors meet
2. Enshrine transboundary object as shared focus for evaluation	2. Multidisciplinary theorization to understand and evaluate object	2. Continuous interaction between evaluator and actors to be evaluated
3. Examine different manifestations of transboundary object	3. Expressive illustration of abstract concepts through tangible examples	3. Ongoing and on-time delivery of findings throughout process

breaking through expectations. Neither is it sufficient to monitor the networking capacity of government agencies without examining their ability to deal with ambiguity, entanglement, and diffusion. In response to the difficulties described above, we felt that our evaluation framework first had to recognize the unpredictable dynamics, unlimited playing field, and contested nature of knowledge. This recognition formed the outer edges of our framework.

We then still had to identify an object to evaluate which would be relevant across the different organizations involved, across the different levels of governments, and across the different disciplines concerned. Drawing on the principles of robust conceptualization, multidisciplinary theorization, and expressive illustration we selected as the central focal point of the evaluation the *intervention capacity* of agencies and stakeholders involved in dealing with terrorism. We defined this as: ‘the combined capacity of the government and its partners to effectively, legitimately, and resiliently contribute towards the goals of the strategy’, in this case preventing terrorist attacks, preventing public fear, and limiting (potential) damage from attacks.

Note that intervention capacity is as transboundary as the terrorist phenomenon it is targeting. Intervention capacity is the composite of all contributing organizations, from National Police to Inland Revenue, but also the personal role of an individual police officer working on the street. National and local, network and personal are combined. The intervention capacity also transcends the different criteria for government action. Only if the intervention capacity is *effective*, *legitimate* and *robust* at the same time, it can be considered as strong overall.

With such a transboundary focus for the evaluation, we still had to find concrete data points or observations to inform our evaluation. We felt that intervention capacity is not easily measured directly; the number of terrorist attacks which do or do not occur will be largely attributable to the context rather than the government capacity. Instead, we focused on the conditions which make it likely for intervention capacity to occur or not. For example, it can be argued that the presence of plans, active cooperation, and strong local initiative are key parts of the collective intervention capacity against terrorists. Of course, additional factors are important as well, such as international cooperation and public communication. However, practical research limitations such as limited time and manpower restricted the scope of our evaluation. These three concepts – plans, cooperation, local interventions – allow us to draw on multiple disciplines to formulate questions, insights, and judgements on the quality of these strategic actions and their anticipated impact on the overall intervention capacity. Our evaluation framework is visualized in Figure 2.

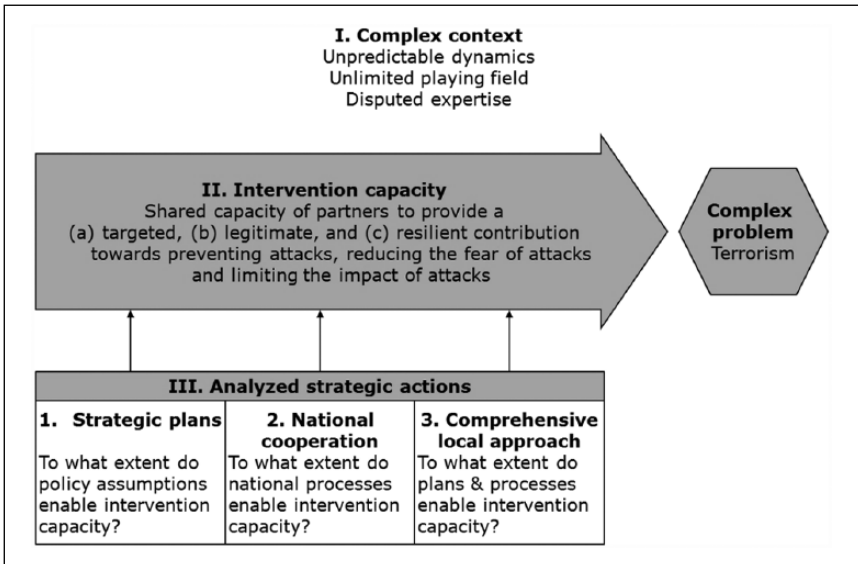


Figure 2. Evaluation focus for evaluating counterterrorism.

Source: Noordegraaf et al. (2016).

Notice that all of the components have clearly tangible objects or locations, but also transcend boundaries in time or place. The plans are a real document, composed in 2011, but the ever evolving assumptions and beliefs about terrorism and counter-terrorism amongst the key players are also taken into account. Similarly, the national cooperation can be symbolized by the Combined Committee for Counterterrorism, but is also present in many bilateral or personal interactions. Finally, the local interventions are small and immediate, but also the manifestations of the wider plans and relationships.

Note that we treated the plans, relationships, and local execution not as corners of a network, stages of the policy process, or layers of the counter-terrorism complex. These parts are all manifestations of the strategy as a whole. In both the plans, national cooperation, and local execution, the same patterns emerged: the debate between prevention versus repression; the division between security versus society-oriented players; the feedback-effects of heavy-handed government interactions. Evaluating the pattern in the small frame allowed use to identify the patterns which shaped the intervention capacity in the larger frame. This reminded us of the relevance of *fractals*, a mathematical concept displaying patterns which repeat themselves across the boundaries of scale.

Evaluation principles to guide evaluators

The formulation of the above evaluation framework and the subsequent assessment of its components was informed by three key principles. These principles emerged at the beginning of our evaluation project and allowed us to connect actors and transgress boundaries between fields of expertise, insights, interests, abilities and acts.

Principle 1: Robust conceptualization

The transboundary nature of counterterrorism made it impossible to ‘reveal’ common ground – but jointly establishing common conceptual ground was possible. We feel this was made possible through the use of robust concepts enabling multiple participants to develop joint views upon the evaluation that is performed, especially on:

- a) *What kind of evaluation study is carried out?* We had to be very precise concerning the nature of our evaluation. This was not only important from a methodological point of view, but also to manage expectations. First, we used the distinction between ‘plan, process, and effect evaluation’, and we stressed the first two elements of these types of evaluation. We also linked these approaches to distinctive evaluation phases. Second, we highlighted the impact of extreme wickedness on the scope of the evaluation, in the sense that causality is contested. On the one hand, we explained the participating organizations that cause-effect links are socially fabricated, for example through communicative processes (fear might result from this). On the other hand, we also explained that causality is circular; actions will not only produce new actions in a given context, but also affect context and original actions.
- b) *Which focus and key terms are guiding the evaluation?* We then had to find the right conceptual and theoretical framing in order to provide a systematic, yet open and inviting focus. The different participating organizations, coming from different disciplines and professional worlds, all had to feel that the evaluation could be relevant and useful to them. As will be discussed in more detail below, we used the notion ‘intervention capacity’ as our focal evaluation framework, highlighting the capability of agencies and stakeholders to intervene in case of terrorism, in ways that are targeted, legitimate and robust. Interestingly, diverse agencies such as the police, intelligence agencies, social workers, and communication specialists all this as a new perspective, but also relevant to their daily work and choices.
- c) *What are the limits of the evaluation?* We had to be realistic as far as evaluation approaches and methods were concerned. The extreme wickedness made it extremely difficult to evaluate actions, let alone effects. Extreme wickedness underscores the importance of stressing the limits of evaluation findings. However, we did emphasize the value of the process. We also positioned our evaluation as a vehicle for bringing agencies and stakeholders together to enhance their joint intervention capacity.

Principle 2: Multidisciplinary theorization

The expansive nature of transboundary issues makes it impossible to evaluate on the basis of rigid and monodisciplinary theoretical models. Problems such as radicalization and terrorism are claimed in equal measure by the disciplines of law, psychology, sociology, history, political science, etc. Privileging one theoretical outlook will limit the knowledge basis of the foundation and undermine the credibility among the disciplines excluded. At the same time, we could not provide a unifying, grand integrative perspective, which cut across the various disciplines and normative frameworks. We therefore tried to work towards theorization bringing together different and even contradictory theoretical elements:

- a) Construct *meta-theoretical outlooks*. We already discussed conceptual framing, generating the need for focused yet open and inviting theory. When we dove deeper into the construction of evaluation theory, we experienced the importance of developing theory that puts things *into perspective*. In our case we used the notion of ‘intervention capacity’ and developed theoretical building blocks to evaluate by relying on existing theoretical outlooks. We argued that the intervention capacity could be manifested in good strategies, collaborations, and local interventions. We could then draw on the relevant literature to evaluate strategy (on the basis of the functions of strategies by Mintzberg 1992 [1987] or planning amidst ambiguity by March and Olsen, 1976), the collaboration in networks (on the basis of the hallmarks of effective collaboration described by Ansell and Gash, 2008), and joint intervention (on the basis of loosely coupled action by Weick, 2005). We so theorized by interrelating and jointly applying several crucial theories coming from different fields.
- b) Create theoretical models that can be *flexibly tied to multiple audiences*. There is an endless number of potential audiences for an extremely complex issues such as counterterrorism. It starts with obvious actors, such as intelligence agencies, the police and criminologists, but it also comprises actors like welfare agencies and housing agencies, as well as schools and teachers. This means there is an endless supply of points of view, which can never be really brought together, let alone integrated. Our theories had to be relevant and valid for the different audiences, we had to be able to translate our concepts and inferences to the multiple audiences involved.
- c) Create *loosely coupled theories*. We relied heavily upon work by Weick (2005), who has not only written about ambiguity and creating institutional capacity to deal with ambiguity; but also about crises and terrorism and the importance of surprise and imagination. His central claim is the importance of ‘looseness’ as opposed to ‘tightness’. In case of ambiguity, crises and surprise, organizations should be about loose couplings as well. Rigidly instructing agencies to always act in a certain way could actually be dangerous in times of crisis. Similarly, we used our theoretical concepts as loosely coupled perspectives. We translated the notion of intervention capacity to various manifestations of the overall strategies, such as the plans, collaborative arrangements, and local interventions. Our yardsticks for determining ‘success’, moreover, were loosely coupled as well. We used multiple yardsticks to assess different elements of intervention capacity, namely *effectiveness*, *legitimacy* and *robustness*, which together represent good governance, but do not always march together.

Principle 3: Expressive illustration

Finally, in order to broadcast our findings to the various agencies, actors and stakeholders, we had to show them what all these theoretical outlook and analytical reflections mean in practice. We could not do this through theories alone, but had to use expressive anecdotes, examples and illustrations.

- a) Use *examples from the field of counterterrorism*. In order to connect theories, methods and data to multiple audiences, we had to show what all these concepts really mean. This started with sketching a timeline of events, facts, and trends which all parties could if not agree with at least engage with. Specific attacks in Paris, Brussels, or other cities provided clear concrete manifestations of the problem at hand. We also used

composite cases of radicalized individuals to concretize what the government response of all the different agencies would look like. The police, intelligence services, telecom operators, and social services all take disparate interventions, but they come together at the level of individual cases of people affected by all of these interventions. We used the example of such cases to concretize our questions and finding.

- b) Use *examples from adjacent or other fields* in order to clarify what they mean. The advantage of using examples from the field of counterterrorism is that connections could be made with technical debates and ‘inner circles’. The disadvantage was that entrenched lines of reasoning were reproduced. Exposing these logics was useful as well, but at times we aimed to forge new connections beyond old divisions. We then used examples from other fields to get the point across. For example, we use the field of medicine to show what intervention capacity means, as well as examples of water safety to show how an effective governance response to wicked issues could look like.
- c) Use *examples to support* key message as well as potential implications. When sketching our final conclusions, it was important to be specific and concrete as well. Not only to elaborate ‘how’ the governance of counterterrorism might be improved, but to connect the perspectives of different actors. In this sense, conclusions and implications were important for *reframing*. Agencies and stakeholders had their own experiences and points of view which would be hard to change. We tried to connect the viewpoints across actors to generate additional intervention capacity. This could make their respective expertise and professional identity less closed and more connective (cf. Noordegraaf, 2015), strengthening their ability for joint action.

Evaluation conditions to engage evaluated actors

Our framework and principles helped us to navigate the conceptual challenges of this evaluation, but we needed some additional guidelines for dealing with the institutional challenges of multiple competing stakeholders. As boundaries between problems, agencies, expertise are confounded in a transboundary issue such as counterterrorism, it is unclear what policy domains fall within the remit of the evaluation, what viewpoints are valid, and even who is being evaluated. The following evaluation conditions helped us to engage and channel the interaction of the organizations and actors being evaluated. These conditions helped to transform the evaluation in one-directional review into more of a shared learning experience actually strengthening the intervention capacity being assessed.

Condition 1: Joint spaces

We need to develop joint spaces during the evaluation to bring the various stakeholders – agencies, experts, organizations, partners – together in order to jointly contribute to the reframing and assessment of the government responses. Most concretely, the advisory board overseeing the evaluation proved to be such a space. The different experts, policy-makers and professionals brought together in this body discussed texts, aligned experiences, and exchanged practices. These policy-makers and professionals, lawyers and sociologists had strongly diverging terminologies, vocabularies, insights, and opinions, but the regular board meetings provided a joint space to manufacture a connected narrative. This narrative had to be clear and convincing for different audiences, as all board members had to subscribe to it, and so became the basis for the language and focus we were able to use in the report and publication of the evaluation to actors further afield.

Condition 2: Continuous interaction

The continuous interaction with the participating organizations was vital to the success of the evaluation. Throughout the evaluation process of a year, we would weekly speak with the different national and local practitioners and experts involved. This was required partly to get to understand counterterrorism domains and to remain informed about current developments, projects and ambitions, but also to facilitate the ongoing alignment of viewpoints in order to generate evaluation outcomes recognized across organizational boundaries. In order to *test, translate and transfer* outlooks, a continuous exchange of views and ideas proved to be necessary. Working in-between terminologies, vocabularies, insights and opinions this exchange could not always produce absolute clarity about events. However, this ongoing interaction did produce legitimacy among actors for the evaluation and acceptance of even the critical findings. To participants, it felt that the evaluation did not produce the exogenous, uninformed judgements of an outsider, but that the evaluation brought together the endogenous, well-informed wisdom of the counterterrorism network itself.

Condition 3: Ongoing and on time delivery

During these continuous interactions in joint spaces, it was vitally important to secure ongoing and on time delivery of the different stages of our evaluation findings. We set and met strict deadlines for sharing intermediate observations or concept conclusions in text form with the various agencies and experts. This meant that momentum was maintained, trust was built and texts could be used to fuel (re)framing. Texts did not so much represent traditional boundary objects, aimed to bring people and perspectives together, but boundary subjects, aimed at more actively involving people and crafting new narratives (see Guba and Lincoln, 1989; Moreau and Clarkin, 2011).

We had to present a clear ‘verdict’ to Parliament, but in the meantime had to align the various agencies and stakeholders so that they would develop common ground and joint views that would enable them to substantiate this verdict *and* link it to practical action so that it would become meaningful to them. As argued, instead of objectively ‘revealing’ the truth about (counter)terrorism, or generating meaningful intersubjective assessments of the phenomenon, we aimed at transgressing boundaries between actors and fields, in such a way that this would contribute to the joint capacities we were evaluating.

Discussion

The response to transboundary wicked issues cannot really be evaluated, but it has to be evaluated. This paradoxical situation has several implications for evaluators, as we experienced in our evaluation of the Dutch counterterrorism strategy. It is essential to stimulate multidisciplinary evaluation. In fact, it is important to turn this into more transdisciplinary evaluation in which new, shared vocabularies arise, which not only bring people together but also enable them to see their joint practices in new light. The term ‘intervention capacity’ proved to perform such a function in our evaluation. Furthermore, in case of issues like terrorism, it is crucial to manage expectations and show things that are understandable for multiple audiences, to clarify what can be expected from which audiences. Finally, it is important to convince others and especially ‘insiders’ by using ‘small’ examples that symbolize and illustrate the ‘big’ picture. Especially, when boundaries are confounded, the bigger picture can be found in small instances, and vice versa. Traditional notions such as ‘macro, meso, micro’ are of little value.

Such instances and examples work as *transboundary objects*, discursively, not so much aimed at bridging gaps and overcoming boundaries but to trace, test and utilize boundaries (cf. Carlile, 2002). In a broader sense, the scholarly work that is put into understanding transboundary wicked issues can be seen as *transboundary work*. It works in the spaces beyond disciplinary, scholarly, practical, societal and (geo)political distinctions. Public governance sciences have the advantage that they add something to the analyses produced by core disciplines like psychology and sociology. They are able to understand plan and processes, for example, in other ways than psychologists and sociologists do. But this does not guarantee common ground and/or shared outlooks.

In addition to relying on symbolic boundary objects such as key concepts and texts, it is important to rely upon other boundary practices that enable evaluators to link disciplines, domains and debates. These practices consist of *objects* such as papers, reports, figures, schemes and boxes, as well as *liaisons*, such as people who already work across fields and in-between disciplines, and *processes*, such as consciously deploying joint platforms. In this way, shared insights are formed and given weight at the same time.

Conclusions

From an evaluation perspective, it is attractive to dive into highly unattractive societal issues. As the societal challenges are intense, evaluators must be prudent and productive at the same time. They need to carefully conceptualize their evaluation frameworks and objects to provide a shared focal point for the evaluation. They must follow clear principles of evaluation which combine robust conceptualization with multidisciplinary theorization and expressive illustration. And they also need to create the appropriate conditions for engaging and channeling the interaction with the organizations being evaluated.

The difficulties evaluators encounter in this work stem from the very nature of the fuzzy, explosive, expansive, and extremely wicked issues they are dealing with. However, their own contribution in evaluation as transboundary work could strengthen the shared ability to address transboundary problems at the same time. An evaluation of one of the greatest threats to society, can also help to reveal one of the greatest benefits of evaluations.

Acknowledgements

The authors would like to thank all the participants of the Dutch counterterrorism evaluation for their insights and contributions, the colleagues at Utrecht University for their multidisciplinary perspectives on evaluation, the participants of the IRSPM 2016 workshop for their feedback, and John Alford for his insightful reflections. They also thank two anonymous *Evaluation* reviewers for their valuable comments.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported was supported by the Dutch WODC, the Scientific Research and Documentation Center of the Ministry of Justice.

Notes

An earlier version of this paper was presented at IRSPM 2016, 13–15 April 2016, Hong Kong, in the panel on ‘Wicked Problems in Public Policy’ (chaired by John Alford and Brian Head). The authors thank the chairs and participants for their valuable comments. A second version was

presented during a research meeting at the Utrecht School of Governance, with another contribution by John Alford.

1. The terms ‘wicked’, ‘wickedness’ etc. are now well-established in both policy analysis and in evaluation. Although this label is acknowledged as such with quotation-marks at first usage, henceforth this is not continued in recognition of widespread acceptance of the term in various literatures.

References

- Ansell C and Gash A (2008) Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory* 18(4): 543–71.
- Ansell C, Boin A and Keller A (2010) Managing transboundary crises: Identifying the building blocks of an effective response system. *Journal of Contingencies and Crisis Management* 18(4): 195–207.
- Baum S and Mitchell WF (2010) People, space and place: A multidimensional analysis of unemployment in metropolitan labour markets. *Geographical Research* 48(1): 13–23.
- Blackman T, Wistow J and Byrne D (2013) Using Qualitative Comparative Analysis to understand complex policy problems. *Evaluation* 19(2): 126–40.
- Boin A (2009) The new world of crises and crisis management: Implications for policymaking and research. *Review of Policy Research* 26(4): 367–77.
- Boin A, Busuioic M and Groenleer M (2014) Building European Union capacity to manage transboundary crises: Network or lead-agency model? *Regulation & Governance* 8(4): 418–36.
- Bore A and Wright N (2009) The wicked and complex in education: Developing a transdisciplinary perspective for policy formulation, implementation and professional practice. *Journal of Education for Teaching* 35(3): 241–56.
- Carlile PR (2002) A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science* 13(4): 442–55.
- Christensen T and Lægreid P (2012) Competing principles of agency organization: The reorganization of a reform. *International Review of Administrative Sciences* 78(4): 579–96.
- Crenshaw M (ed.) (2010) *The Consequences of Counterterrorism*. New York: Russell Sage Foundation.
- De Graaf BA (2010) *Theater van de angst. De strijd tegen terrorisme in Nederland, Duitsland, Italië en Amerika* [Theatre of Fear. The Battle against Terrorism in The Netherlands, Germany, Italy, and America]. Amsterdam: Boom.
- Djelic ML and Sahlin-Andersson K (eds) (2006) *Transnational Governance: Institutional Dynamics of Regulation*. Cambridge: Cambridge University Press.
- Ferlie E, Fitzgerald L, McGivern G, et al. (2011) Public policy networks and ‘wicked problems’: A nascent solution? *Public Administration* 89(2): 307–24.
- Guba EG and Lincoln YS (1989) *Fourth Generation Evaluation*. London: SAGE.
- Hannigan B and Coffey M (2011) Where the wicked problems are: The case of mental health. *Health Policy* 101(3): 220–7.
- Harris J, Brown VA and Russell J (eds) (2010) *Tackling Wicked Problems: Through the Transdisciplinary Imagination*. London: Earthscan.
- Haynes P (2007) Chaos, complexity and transformations in social care policy in England. *Public Money and Management* 27(3): 199–206.
- Head BW (2008) Assessing network-based collaborations: Effectiveness for whom? *Public Management Review* 10(6): 733–49.
- Head BW and Alford J (2015) Wicked problems implications for public policy and management. *Administration & Society* 47(6): 711–39.
- Koppenjan JFM and Klijn EH (2004) *Managing Uncertainties in Networks: A Network Approach to Problem Solving and Decision Making*. London: Routledge.
- Koppenjan JFM and Klijn EH (2013) What can governance theory learn from complexity theory? *Mirroring two perspectives*. In: Keast R, Mandell MP and Agranoff R (eds) *Network Theory in the Public Sector: Building New Theoretical Frameworks*. New York: Routledge, 187–206.

- Lawrence R (2011) Deriving collaborative aims and outcomes: A case-study of cross-border cooperation in Central and Eastern Europe. *Evaluation* 17(4): 365–82.
- Levin K, Cashore B, Bernstein S, et al. (2012) Overcoming the tragedy of super wicked problems: Constraining our future selves to ameliorate global climate change. *Policy Sciences* 45(2): 123–52.
- Lum C, Kennedy LW and Sherley A (2006) Are counter-terrorism strategies effective? The results of the Campbell systematic review on counter-terrorism evaluation research. *Journal of Experimental Criminology* 2(4): 489–516.
- March JG and Olsen JP (1976) *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget [University Press].
- Mintzberg H (1992 [1987]) Five Ps for strategy. In: Mintzberg H and Quinn JB (eds) *The Strategy Process*. Englewood Cliffs, NJ: Prentice-Hall.
- Moreau KA and Clarkin CL (2011) Thinking beyond measurement, description and judgement: Fourth generation evaluation in family-centered pediatric healthcare organizations. *Journal of MultiDisciplinary Evaluation* 8(17): 33–45.
- Noordegraaf M (2015) *Public Management. Performance, Professionalism and Politics*. Basingstoke: Palgrave.
- Noordegraaf M, Douglas S, Bos A, et al. (2016) *Evaluation Counterterrorism Strategy 2011–2015*. Utrecht: Utrecht University - School of Governance (USGO), WODC.
- Rittel HW and Webber MM (1973) Dilemmas in a general theory of planning. *Policy Sciences* 4(2): 155–69.
- Schmid AP (2013) Radicalisation, de-radicalisation, counter-radicalisation : A conceptual discussion and literature review. *ICCT Research Paper*. Den Haag.
- Termeer CJ, Dewulf A, Breeman G, et al. (2013) Governance capabilities for dealing wisely with wicked problems. *Administration & Society* 47(6): 680–710.
- Turnpenny J, Lorenzoni I and Jones M (2009) Noisy and definitely not normal: Responding to wicked issues in the environment, energy and health. *Environmental Science & Policy* 12(3): 347–58.
- Van Bueren EM, Klijn EH and Koppenjan JF (2003) Dealing with wicked problems in networks: Analyzing an environmental debate from a network perspective. *Journal of Public Administration Research and Theory* 13(2): 193–212.
- Walton M (2016) Expert views on applying complexity theory in evaluation: Opportunities and barriers. *Evaluation* 22(4): 410–23.
- Weber EP and Khademanian AM (2008) Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Public Administration Review* 68(2): 334–49.
- Weick KE (2005) Organizing and failures of imagination. *International Public Management Journal* 8(3): 425–38.
- Williams CR and Arrigo BA (2002) *Law, Psychology, and Justice: Chaos Theory and the New Order (Dis)*. New York: SUNY Press.

Professor Mirko Noordegraaf is Full Professor of Public management at the Utrecht University School of Governance (USG), The Netherlands.

Dr Scott Douglas is Assistant Professor of Public management at the Utrecht University School of Governance (USG), The Netherlands.

Aline Bos, MSc, is Lecturer and Academic Consultant at the Utrecht University School of Governance (USG), The Netherlands.

Wouter Klem, MA, is PhD candidate at the department of History at Utrecht University (ERC funded project ‘Securing Europe, fighting its enemies 1815–1914’).