



UvA-DARE (Digital Academic Repository)

Cutting through complexity

Evaluating countering violent extremism (CVE)

Gielen, A.-J.E.

Publication date

2020

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Gielen, A.-J.E. (2020). *Cutting through complexity: Evaluating countering violent extremism (CVE)*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

1 Introduction

1.1 The complexity of violent extremism and challenges for evaluation

Amira¹ is a 16 year old schoolgirl in a small town in the Netherlands. She has lived with her mom and sisters since her parents got divorced. The family has a long history of engagement with child protection services because of repeated domestic violence by the father, often instigated by his alcoholism. Amira's parents are refugees and her mother is a Muslim. Amira herself is not concerned with Islam. She is a typical teenager. She hangs out with friends and is absorbed with clothes, make-up, boys and social media. That all changes when her mother is diagnosed with cancer and nearly dies. As a divorced refugee, her mother has very little by way of social or family network she can draw on while she is so sick. Word goes round in the local Muslim community and young women with a Salafi religious orientation start helping out. They also start speaking with Amira, observing that she is not behaving like a 'true' Muslim. For example, she does not wear a head scarf. These women introduce her to books, websites and lectures so she can learn more about Islam and how she should behave and act as a 'true' Muslim. After the summer holiday Amira makes a radical decision. She decides to swap her skinny jeans for a khimar; she is completely veiled in black, with the exception of her eyes, nose and mouth. This choice generates a lot of resistance, both at school and in town. She is nearly run over by a driver who shouts 'terrorists' at her; and her maths teacher dedicates his lessons to her new choice of clothing, continually asking her if she is going to Syria (to join ISIS). At that point 'going to Syria' is not on her radar at all, but her new clothing and religious quest is attracting attention on social media. Amira even receives a marriage proposal from a 'mujahid' (a fighter in Syria). It turns out that an old friend of one of her sisters has already travelled to Syria. She becomes more and more engaged in the extremist network, both via social media as well as via other women she knows locally. She becomes part of a pro-ISIS What's App group in which members share thoughts about western 'hypocrites' and 'hijra', their Islamic duty to migrate to an Islamic country, the so-called caliphate. Joining this group gives Amira a sense of belonging and recognition: at least she is not the only one who feels discriminated and feels she might be better off in the caliphate. Through a network in Syria, Belgium and the Netherlands, money

1 To ensure privacy, the name and age of Amira have been changed. See also section 1.6

is transferred into her account, and the following week she decides to travel to ISIS, but gets stopped at the Turkish border.

Peter² is a 17 year old high school student and very much engaged in politics. His parents are still together, giving him a very protected and strict upbringing. He is never allowed to go anywhere or to do normal teenage things. This becomes an issue when he starts secondary school in a big city. It is one of the reasons why he gets bullied at school. He is particularly bullied by children with a different ethnic background. Combined with the fact that he is very much engaged in right-wing politics, the experience of bullying fuels his sense that 'foreigners' are bad and that the idea of a 'multicultural society' is a failure. In his online searches on the failure of the multicultural society he meets John on an internet forum. They start chatting online about their similar political ideas. Soon John invites him for drinks at an organization called 'Voorpost'.³ During these events 'comrades' discuss different societal issues and their often extremist right-wing views on these. For Peter these events are like 'coming home'. Finally he has a group of mates who think like him and with whom he can voice his political opinions without being frowned upon. He is no longer ashamed to say what he thinks and to stand up for his beliefs. In fact, Peter decides to shave his head and wear army boots to school. He even hands in an essay at school in which he provides fifty arguments for why the Holocaust could never have taken place. His history teacher is furious. How dare a student write such nonsense! For Peter this is just confirmation that the multicultural society stifles his right to freedom of expression. He becomes even more set in his views, and starts participating in demonstrations organized by Voorpost. After one of these demonstrations, Antifascistic Action (AFA), an extremist left-wing organization, attacks him and his friends. They are also named and shamed on the AFA website as nazis. This calls for revenge. When they learn that the AFA has squatted a building in their city they make plans to smoke them out. One evening they storm the building wearing balaclavas and armed with an extendable baton, pepper spray, an air gun and Molotov cocktails. The AFA squatters manage to escape, but the building is set afire. In the weeks that follow, Peter and his friends are arrested for arson.

2 *Ibid.* (see also Gielen, 2008). See also section 1.6

3 Voorpost considers itself a nationalistic action group. They oppose the EU, European cooperation, the Belgian state and are proponents of an independent Flanders and a reunited Netherlands. Further they are anti islamization, globalization and Americanism, calling for cultural survival.

The above cases are two real life stories of different extremists I have worked with and interviewed in my decade as a researcher, advisor and intervention provider in the field often designated as ‘countering violent extremism’ (henceforth, CVE⁴). The stories depict completely different cases in which two individuals become engaged in different forms of violent extremism: jihadism and right-wing extremism. While the root causes and pathways toward extremism differ, the stories do exhibit one similar and important aspect: the complexity of a radicalization process. This complexity was an issue that I addressed more than a decade ago in my master’s thesis.

In the aftermath of the 2004 murder of Dutch filmmaker Theo van Gogh ‘counter-radicalization policies’ were initiated in the Netherlands. As a political science student, I was intrigued by the debate on the question of what government could effectively and legitimately do to prevent and counter violent extremism. I started ‘evaluating’ counter-radicalization policies in Amsterdam boroughs as part of both my bachelor’s thesis (Gielen, 2006) and my master’s thesis (Gielen, 2007), which were eventually published (Gielen, 2008; Uitermark & Gielen, 2010). I soon reached the conclusion that I could not evaluate these policies properly without gaining a better understanding of why and how people radicalize. After all, how could I reliably assess whether counter-radicalization policies were effective, if I did not have a well-grounded understanding of radicalization and its causes, or of the accuracy of designating the target audiences of these governmental policies as ‘radical’ or ‘extremist’? Also, I wondered how I could determine whether these groups were indeed less vulnerable to radicalization and violent extremism after being exposed to interventions to counter violent extremism?

I tried to overcome these research challenges in two ways. First, I immersed myself in groups of radicalizing youths with Muslim and right-wing extremist backgrounds. Encounters during this period of ethnographic work brought me stories like the ones summarized in the above vignettes (Gielen, 2008). They made me deeply aware of the complexity of radicalization as a phenomenon. Each individual’s potential path toward radicalization is shaped by manifold factors, interacting in multiple ways. These factors, as well as the way a person deals with them, differ between individuals and across contexts. Second, I decided to embark on a literature search, looking for theories and models to explain how and why people start to adopt views that legitimate violent action, and that may lead them to join a radical organization or even to commit acts of violence (for now, I refer to this process as

4 This concept is further discussed in section 1.3

'radicalization'). During my quest, I soon noticed that this complexity was not in fact adequately addressed in the existing literature. To be sure, many of the factors and interactions I had observed were covered, but virtually none of the theories covered the whole range. Furthermore, and at least as important, the ways such factors, their interactions, individuals' agency and context worked together were undertheorized, to put it mildly. I found that, depending on their own scientific discipline, scholars drew on relative deprivation theory (Gurr, 1970; Buijs *et al.*, 2006; Veldhuis & Staun, 2009), social movement theory (Wiktorowicz, 2004), network theory (Sageman, 2004), theories from social psychology (Kruglanski & Webber, 2002) and identity theory (Tajfel & Turner 1986; Buijs *et al.*, 2006; Schwartz *et al.*, 2009). Some of these theories, and combinations thereof, led to development of sequential models for the process of radicalization, such as Moghaddam's (2005) 'staircase to terrorism' model and McCauley & Moskaleiko's (2008) 'pyramid' model.

Based on my ethnographic work, I realized that any attempt to understand radicalization and evaluate measures to prevent and counter radicalization, would have to account for the diversity and complexity found. I thus decided to embark on another adventure: developing a new evaluation approach. Thus equipped, following my bachelor's and master's studies (Gielen, 2007, 2008; Uitermark & Gielen, 2010), I started my own, independent consultancy specializing in evaluation of what was then in the Netherlands called 'anti-radicalization measures'. It did not take long to realize that the so-called 'gold standard' (Timmermans & Berg, 2003) for evidence-based work, meta-analysis based on randomized controlled trials (RCTs), would be of no real help here. As had also become clear in cases of non-communicable diseases (see, e.g., Moret *et al.*, 2007; Gielen & Grin, 2010), evaluations based on RCTs offered severely limited potential to understand interventions in complex phenomena, be they irritable bowel syndrome, lower back pain or radicalization. This is due to their narrow focus on linear causal relations between a limited set of values, neglecting agency and contextuality.

Obviously, a new evaluation approach was needed; one that recognized complexity, including the role of context and agency. Ideally, this new approach would incorporate what we already understand as a basis for evaluating measures, and also enable evaluation results to be fed back to refine that understanding. Such a method would help address a key methodical gap, contribute to knowledge accumulation and provide part of a basis for further theory development in the field. I therefore started

to explore modes of evaluation that would not answer the question of *if* an intervention worked (yes or no), but rather provided insight into *how* interventions might work *in a particular context*. Such insights have much more to offer for the design and redesign of interventions and theory building than simple yes/no answers on effectiveness questions.

It is this adventure that I report in the current dissertation. While my underlying personal interest and the academic knowledge gaps will, by now, be roughly clear, it is high time to elucidate the societal relevance of my efforts. In a sentence: All too often policy measures and societal discussions, much like the academic literature, fail to take complexity into account, and current evaluation approaches tend to reproduce rather than remedy this failure, though all this has more than merely academic consequences.

1.2 Policymaking and societal debate: The repercussions of neglecting complexity

In order to appreciate the societal relevance of this dissertation, it is useful to start with two example cases. These are cases that indicate how and why lack of full recognition of the complexity of radicalization may have consequences for policymaking and proper judgement in the public sphere.

In France, Dounia Bouzar and her Centre for the Prevention of Sectarian Trends (*Centre de prevention des dérives sectaires liées à l'Islam*) became a trusted partner of the French government and was contracted to implement as part of the government CVE programme. Bouzar claimed that she had deradicalized many ISIS recruits (mostly women). This attracted media attention from around the globe (Symons, 2016). Her methods focused on disengaging the sectarian grip (Bouzar *et al.*, 2014; Bouzar, 2017). This was done via 'de-indoctrination sessions' with parents and radicalized individuals. Parents were either individually coached or supported in family group sessions to help them re-engage with their children and reawaken their emotions. This was done, for example, by singing a song or reminiscing about a specific holiday (Bouzar *et al.*, 2014: 30). Radicalized individuals had to attend meetings at which former extremists or families of a former extremist, provided testimonials.

While working with parents is considered an important element in exit programmes (Gielen, 2018a; Sieckelink & De Winter, 2015; RAN, 2019), it is

just one of several elements that need to be addressed. Based on scientific insights on both radicalization and deradicalization, we know that for success, exit programmes require a holistic approach that addresses the multiple causes of why an individual has become radicalized (McDonald, 2018; RAN, 2019; in line with my earlier ethnographic work, cited above). Yet, personal characteristics, mental health, trigger factors, and micro and macro environmental issues all seem to be bypassed in Bouzar's approach to disengagement. The French government has meanwhile ended its working relationship with Bouzar and her centre (Uhlmann, 2016), and the French approach has shifted toward setting up deradicalization centres. Of the 13 planned centres, however, only one was realized. It opened its doors in September 2016, only to be closed less than a year later, following a report published by the French senate criticizing the deradicalization programme on offer. The French Ministry of Interior released a press statement characterizing the centre as an exercise in 'limitations' (*Le Monde*, 2017). The centre, which worked on the basis of voluntary enrolments, 'de-indoctrination' (which is not the same as addressing ideology and religion) and vocational training, had only nine participants, none of whom completed the programme (Uhlmann, 2016; Sage, 2017).

Similar to Bouzar's approach, the centres' programme did not address the multiple factors that play a role in the radicalization process, and this had consequences for the effectiveness of the policy. This experience raises the question of whether exit programmes can be effective. Or, to be successful, what specific underlying conditions and mechanisms are required? Must we conclude on the basis of the French experience that voluntary exit is always ineffective? Or would merely the application of different interventions do the trick? Should interventions address or avoid the ideological aspect of radicalization? The French example underlines the need for more evaluation of CVE policies and programmes.

A similar case in which the complexity of radicalization was ignored in policymaking was in Amsterdam. Here, Mayor Eberhard van der Laan largely disregarded the ideological dimension of violent extremism (Gielen, 2017c). Rather, the views of one of the mayor's foremost advisors, David Kenning, became leading in the city. Kenning argued that with the new generation of jihadists, ideology was not an important factor. Rather, radicalization toward violent extremism was a consequence of issues related to criminality, family, identity, anger and islamophobia (Van Teeffelen, 2017). Thus Amsterdam's CVE policy focused on dealing with the socio-economic and

cultural breeding grounds for radicalization, neglecting other important causes of radicalization (Gielen, 2017c).

In each of these two cases, the story started with an understanding of the problem of radicalization which, by drawing on a too limited or even flawed set of ideas on radicalization, neglected much of the issue's complexity. However, the problem definition shapes the search space for solutions (Schön, 1983; Hoppe, 2011). Thus, solutions emerged that also failed to sufficiently take into account the complexity of the radicalization processes they sought to address. As these stories show, this may lead to ineffective policies and heated debate. Thus, radicalization is not, or rather not effectively, addressed, and public debate produces more heat than light, threatening the reputations and positions of key players. This is not the place to analyse at length why problem definitions are so limited. For our purposes here it suffices to say that the reasons encompass personal beliefs, ideologies, worldviews and institutions. How this may work became clear in a recent study by Warrington (2018) on Danish governmental policies targeting online radicalization. Through a discourse analysis of policy documents and open, semi-structured interviews with policymakers, she identified six 'logics' behind assumptions regarding the dangers of online radicalization that shaped solutions in specific ways. Interestingly, these assumptions based on partial, sometimes intuitive, understandings⁵ ultimately interfered with two perhaps more valuable logics. The first is the very recognition of online radicalization as a new and complex phenomenon about which 'we' know little, but which does pose an urgent risk – this latter conviction being reinforced by the securitization discourse prevailing in Denmark. The second is the awareness that the lack of solid, reliable evaluation methods leaves no other option than to just try to keep what appears to be working.

Let us reflect on these realities from a policy science perspective. If anything, these examples exhibit a seriously underdeveloped process of what Hoppe (2011: 27–31) called 'problem structuring'. Problem structuring implies two,

5 Including the assumption that online radicalization is a "gateway to offline engagement" (Warrington, 2018: 134), the basic sense that "*sometimes you just have to try something*"; the assumption that an inclusive, online debate with engagement of civil society would help overcome issues of legitimacy and trust in the state; and the assumption that extremist content cannot be eliminated suggests the need to promote "critical thinking" (*ibid.*: 137) as a solution. This is further confirmed by the (normatively induced) assumption that socio-economically marginalized youths may be especially susceptible to such content. Thus arise solutions relating to counter-narratives, online discussion fora, etc.

interrelated and iterative, activities: problem finding and problem solving. In proper, inclusive problem finding one goes back and forth between the conditions underlying a problem, and the diverse ways these are portrayed in various actors' problem representations, so as to find a problem definition that sufficiently covers all the various conditions and representations.

In our examples, as well as in the Danish case, problem finding was far from ideal: the set of problem representations and problem conditions scrutinized for problem finding appears to have been rather limited. Even without in-depth analysis, it is easy to recognize aspects from the literature on how this might arise from the interplay between normative assumptions (more or less intuitive empirical preconceptions) (Dunn, 2015 [1980]; Webber, 1991) and more developed elements constituting 'policy theory' (Hoogerwerf, 1990; Bressers & Hoogerwerf, 1995) or a 'theory of change' (Weiss, 1995). As Hoppe (2011: 71) emphasized, this may be further directed by organizational and individual heuristics ('bounded rationality'), seeking a 'fit' with the outside world ('ecological rationality') or considerations to which one feels accountable ('social rationality'). The second activity of problem structuring is problem solving; that is, the construction, application and evaluation of solutions that may effectively and legitimately⁶ address the problem. Problem solving iterates with – informs and is informed by – the problem finding. This iteration underlies the contingent relationship between policy problems and solutions, as found by Peters (2005: 351) in an insightful overview of the policy literature. It yields what Teisman (1993) and Hoppe (2011) referred to as 'problem-solution couplings'. Thus, the limited problem finding witnessed in our cases may lead to limited solutions that bypass key problem conditions and representations. If policy work becomes focused on such solutions, blind spots in problem finding may be further reinforced.

These difficulties are aggravated by the fact that radicalization is far from a simple, structured problem. In addition to the inherent complexity discussed above, which leaves our empirical understanding limited at best, actors involved have different and often competing beliefs, values and worldviews. These, in turn, produce a plurality of policies and practices that are not necessarily aligned with common scientific insights. In the words of Rob Hoppe (2018: 389) "what one 'sees', deems 'relevant facts', accepts as 'evidence',

6 Of course, a solution may derive its legitimacy partly from the very fact that it addresses a problem. But this may not be enough: the solution may have features disliked by some groups in society, for instance. Thus problem solving is as much a political process as problem finding

and as ‘potential solutions to a problem’ may widely differ”. These features together make radicalization a so-called wicked problem (Ritter & Webber, 1973; Hayden, 2006; Camillus, 2008; Mertens, 2015; Fischbacher-Smith, 2016; Noordegraaf *et al.*, 2017). In cases of wicked problems, problem structuring must essentially be based on an intertwinement of political discussion and analysis. A pertinent question for evaluators of measures in our field is how their work may contribute to that process.

While there is an abundance of literature on how to move from a complex and wicked problem toward a more tamed and structured one (Schön & Rein, 1994; Hoppe, 2011; 2018, Hayer & Laws, 2006), very little has been written on how to evaluate complex and wicked problems. Some have proposed deliberative or interactive forms of evaluation (see, e.g., Guba & Lincoln, 1989; Grin & Van de Graaf, 1996; Loeber, 2004), in which actors with different normative and empirical perspectives discuss the effectiveness of policy measures. Evaluation then becomes a form of political judgement, in which views on what is effective and desirable inform and correct each other. The point is, however, that this requires some basic common ground, such as at least a rough delineation of the problem (Grin *et al.*, 1997: 39; Loeber 2004: 291–293). Yet, the examples above cast doubt on whether a field as complex as countering radicalization can in fact be evaluated, as the actors, factors, conditions and contexts to be taken into account are hard to delineate. Indeed, Noordegraaf *et al.* (2017: 393) questioned whether something as complex as measures to prevent and counter radicalization can be evaluated at all. They considered it to be a *transboundary wicked problem*, which “transcends the distinction between cause and effect, local and global problems and facts and impressions” (*ibid.*: 289). These authors therefore proposed a transboundary evaluation approach consisting of three components. The first component of the evaluation framework is ‘robust conceptualization’, entailing that stakeholders develop a shared vision of the type, focus and limitations of the evaluation. The second step is multidisciplinary theorization, in which a theoretical model is developed based on different scientific disciplines. Finally, expressive illustration is required, in the form of anecdotes, images and examples, to disseminate the findings to the different stakeholders.

Noordegraaf *et al.* (*ibid.*) developed and applied this transboundary evaluation framework to evaluate the Dutch national counterterrorism strategy 2011-2015. This led them to shift their focus from evaluating interventions to ‘intervention capacity’, defined as “the combined capacity of agencies

and stakeholders involved in dealing with terrorism” (*ibid.*: 398). While this is useful in and of itself, it also shifts the burden of knowing what to do to the practitioners making up that capacity and inhibits proper, independent evaluation of their efforts. In addition, important ambiguities remain as to how to decide, for instance, what stakeholders should be selected, for example, to develop a shared evaluation focus, and why. Which theories from which disciplines should be included and excluded in development of the multidisciplinary theoretical model? This framework is essentially a more tailored form of the interactive evaluation approach of Guba and Lincoln (1989), which led to a shared thick description of the Dutch approach to counterterrorism. However, this exercise provided no insight into the degree of success of Dutch counterterrorism policy measures and the relevant contextual conditions. Neither did it offer opportunities to learn from similar policies in other countries. For precisely this reason, Guba and Lincoln’s (*ibid.*) evaluation approach has been criticized (Pawson & Tilley, 1997: 23).

In fact, this mutual reinforcement between lack of proper theory and lack of good evaluations was a key motive for me to embark on my quest for modes of doing evaluation that also contribute to understanding *how* measures do or not produce outcomes. The reflections above help to further sharpen the mission of that journey: we may use evaluations not only to understand *what has been left out* of consideration in a particular policy process or societal debate, but also to learn *how diversity and agency might matter* in producing outcomes. In other words, by casting a wider net and remaining open to additional factors and interpretations, evaluations may achieve two key objectives:

- unmasking too simplistic solutions and their associated problem definitions;
- finding out how factors *not* originally taken into account in an intervention and in the associated debate play a role in the working-in-context of that intervention, thus enriching not only practice with an enlarged understanding, but also correcting problem definitions based too heavily on the views of certain actors rather than the problem at hand.

Such forms of evaluation could contribute to policymaking and debate in interesting ways. Not only would they demonstrate the weaknesses of specific, too simplistic, limited interventions, even more importantly, they would contribute to better problem finding in policy practice, while advancing societal debate. In the longer term, they would further develop the knowledge bases that inform problem finding and solving.

In this dissertation, I discuss my attempts to further develop, from this perspective, realistic evaluation as an input that may bring policymaking and societal debate to a qualitatively higher level of understanding. Specifically, realistic evaluation may help make policies more effective. At least as important, while it is obviously not in itself sufficient to bring normative debate to a conclusion, it might help to clarify the merits, limits and blind spots of different problem representations, as well as their mutual relations. Before I discuss realistic evaluation and realistic review further, however, we must first look at CVE policies more generally. What kind of policy programmes have been developed in Europe? What are the origins of these policy programmes? What are their main elements? This information provides a starting point for understanding the relevance and potential of the realistic evaluation approach, but also its limits and current flaws.

1.3 CVE as an essentially contestable concept

One analytical implication of the fact that radicalization is a wicked problem is that even its definition, as a phenomenon, is complicated and without consensus. While anti-radicalization policies are widely accepted in the Netherlands and other European countries, in scholarly literature and policymaking the term countering violent extremism (CVE) is more common, as discussed widely and critically. In line with this critical debate, this dissertation defines CVE as follows:

Proactive actions to counter efforts by extremists to recruit, radicalize, and mobilize followers to violence. Fundamentally, CVE actions intend to address the conditions and reduce the factors that most likely contribute to recruitment and radicalization by violent extremists. Where possible, CVE should be incorporated into existing programs related to public safety, resilience, inclusion, and violence prevention. CVE efforts do not include gathering intelligence or performing investigations for the purpose of criminal prosecution. (U.S. Department of Homeland Security, 2019)

After the 2004 Madrid bombings and attack on filmmaker Theo van Gogh in the Netherlands and the London '7/7 attacks' in 2005, various EU member states adopted preventative approaches to radicalization and violent extremism, also known as 'counter-polarization and radicalization' programmes, or preventing violent extremism and deradicalization. These approaches are similar in essence (Bakker, 2015; El Difraoui *et al.*, 2015; Lindekilde, 2012).

They attempt to address (i) the root causes and drivers of radicalization and violent extremism and to enhance protective factors against violent extremism, (ii) they understand radicalization as a process in which violent extremism is the final stage of an individual trajectory in which world views and ideology become increasingly more 'extreme', and (iii) these programmes target specific categories of people. This section briefly discusses these main elements of CVE policies in Europe.

First, while there is no academic consensus on what precisely causes violent extremism, there is now general agreement among CVE researchers and policymakers that an interplay of different grievances, causes and risk factors plays a role in violent extremism. Borum (2015) identified so-called risk clusters related to violent extremism: affections/emotions, behaviours, cognitive style, beliefs/ideology, attitudes, social factors, identities and capacities. Similarly, Ranstorp (2016) offered a 'kaleidoscopic overview' of nine risk factors related to violent extremism: (i) individual and (ii) social factors such as frustration and alienation; (iii) political and (iv) ideological factors such as dissatisfaction with foreign policies and interference with religious practices; (v) identity crises reinforced by a migration biography or post-traumatic stress disorder; and (vii) group dynamics and recruitment strategies consisting of (viii) groomers and (ix) social media (Gielen, 2017b). Campelo *et al.* (2018) conducted a systematic literature review of causes of violent extremism in Europe. That review highlights a multitude of risk factors at the personal, micro-environmental and macro-environmental level. At the individual level, the study identifies different traits and psychological vulnerabilities, such as having a depressive nature; experiences with addictive and risky behaviour; early experiences of abandonment; vulnerable family structures; a quest for an ideal; personal uncertainty; perceived injustice; trigger events such as trauma and (near) death; and psychopathological mechanisms that reinforce extremist engagement. Micro-environmental factors may include having radicalized peers, a dysfunctional family and sympathy with the sectarian use of dehumanization to justify violence. At the macro-environmental level they identified the following risk factors: social polarization, perceived group threat, the role of religious ideology, the geopolitical context and processes of societal change. However, since radicalization is so complex and the identified root causes are so numerous it is impossible for any policy programme to cover all the factors and causes, in addition to the new explanatory factors and root causes that are still to be discovered, and thus also need to be included.

Second, the CVE approach is primarily based on the normative idea that violent extremism should be understood as a *process*. This in turn suggests that violent extremism is governable through risk assessment and preventive interventions in the early stages of this process (Githens-Mazer, 2012). Heath-Kelly (2013) argued, for instance, that the concept of radicalization enables policymakers to formulate a linear narrative around violent extremism. This makes the narrative accessible to problem-solving approaches. As such, intervening in the 'radicalization process' is understood as a crucial component of countering violent extremism, through a focus on categories of risk. Kundnani (2012: 4) argued that the concept of 'radicalization' emerged as a vehicle for policymakers to explore the process of how someone becomes a violent extremist and to provide an analytical basis for preventative policies that go beyond the threat of violence or detention. According to him, radicalization models encourage authorities to believe that they can prevent future terrorist attacks by intensive surveillance of Muslims in Europe. The purpose of such surveillance is to establish whether these individuals express world views that can be defined as extreme – though it must be recognized that no convincing demonstration has been provided of any causal relationship between such world views and violence.

The third important element of CVE policy in Europe is that it is aimed at various but specific target audiences: communities or religious congregations, particular groups (e.g., youth groups and schools) and vulnerable individuals. To reach this varied population, countries across Europe have implemented CVE policies consisting of elements like awareness raising and training for practitioners; community engagement; education of young people to make them more resilient to the draws of violent extremism; multi-agency approaches to signal and assess vulnerable individuals; and deradicalization and disengagement programmes for violent extremists (Bakker, 2015; De Goede & Simon, 2012; Gielen, 2017b; El Difraoui *et al.*, 2015; European Commission, 2016; Van Ginkel & Entenmann, 2016; Wittendorp *et al.*, 2017).

Yet, CVE and these underlying assumptions are contested, especially by scholars in terrorism studies who apply critical theory, such as Foucault's post-structuralism and Althusser's structuralism. In essence, their critique is that this field is dominated by a common set of narratives, beliefs and assumptions on (countering) violent extremism which is often informed by social, political and personal beliefs that are continually reproduced and empirically unchallenged (Jackson *et al.*, 2007; Heath-Kelly, 2013, Kundnani,

2012, 2015). Heath-Kelly (2013: 412) reviewing the policy programmes of the UK, concluded that “as post-structuralist scholars have argued about projects of surveillance and screening, such security practices may act in the name of managing risk but they are actually performative of it – producing risk to enable interventions”. Using such a critical theoretical perspective, scholars have observed that CVE programmes primarily legitimize anticipation of security threats to enable intervention at the earliest possible stage, drawing on the idea that radicalization is ultimately (i) a linear process (ii) in which individuals enter into a radicalization process and (iii) then develop and display behaviour or express particular ideas that can be identified as part of a radical belief system (De Goede & Simon, 2012). CVE programmes have the goal of identifying these behaviours and ideas, and intervening and changing them.

I agree that by focusing on specific root causes, by viewing radicalization as a process and by targeting specific communities, CVE policies risk being based on flawed or normative assumptions and eventually may even produce counter-productive effects. Defining and identifying radical behaviour is always relational and normative, following from the ideas, values or conduct that the radicalized individual is being compared to (Vermeulen, 2014). In terms of root causes, there is always a normative dimension in policymakers’ identification and addressing of particular root causes.

As I have argued, radicalization is a complex and wicked problem with many different explanatory factors. First, identifying and intervening in particular root causes will always ignore other root causes, which will affect the outcomes of the chosen policy programmes. In evaluating these programmes these decisions need to be taken into account. Second, viewing radicalization as a process evokes the idea among policymakers that they can intervene in the process, starting with identifying individuals who are in the early phases of radicalization. The focus on risky individuals then often becomes a focus on ‘people at risk of becoming risky’. The explicit link with violent extremism becomes more and more blurred (Heath-Kelly, 2013). Third, a critical examination of the UK’s Prevent policy, in which the government sought to prevent and counter violent extremism by engaging with Muslim communities, has provided insight into the associated risk of stigmatizing Muslims and creating ‘suspect communities’ (Kundnani, 2009, 2012; Bigo *et al.*, 2014; Vermeulen, 2014).

We have thus established that violent extremism is a wicked problem. CVE is also a contested concept. It is criticized as being based on normative

choices on which root causes are and are not to be addressed. Moreover, radicalization is typically depicted as a linear process that can and should be countered at the earliest possible stage. Furthermore, the choice of which specific audiences are to be targeted in the implementation of CVE policies is a normative one. These issues need to be addressed when evaluating CVE.

1.4 Realistic evaluation and realist review

CVE policy is essentially designed and implemented as a 'battle of ideas' revolving around this complex and wicked problem. This poses challenges for evaluation. However, more than a decade after the conclusion that "CVE evaluation is still in its infancy" (Nelen *et al.*, 2010), CVE evaluations still remain scarce and CVE evaluation practice continues to be underdeveloped (Bellasio *et al.*, 2018). John Horgan, one of the world's leading terrorism experts, stated the following on the lack of CVE evaluations:

One thing that many of these [CVE] programmes have in common, however, is their resistance to evaluation. I am in favour of innovative and evidence-led CVE programming, but programs that shy away from evaluation offer little more than smoke and mirrors. Evaluation needs to happen early and often. For the most part, it remains an afterthought. A CVE programme without an evaluation component is, in my view, worthless (European Eye on Radicalization, 2018).

The dearth of CVE evaluations is due to several factors related to the wickedness of (countering) violent extremism: (i) the lack of consensus on how the many root causes of and factors in violent extremism interplay and thus what combination of interventions is essential to successfully counter violent extremism; (ii) the broad array and complexity of interventions across the CVE spectrum; (iii) the lack of indicators and tools for measuring the outcomes of prevention; and (iv) the prominent role of context when implementing CVE, as for example, political context, timing and geography all influence CVE effectiveness (Gielen, 2017b). CVE evaluations could, however, be of academic and policy relevance for several reasons. CVE evaluation can enhance the legitimacy of this often contested policy by providing clarity on the (normative) assumptions that CVE policies are based upon, thus unmasking too simplistic solutions and contingent problem definitions. The magnitude and growth of the 'deradicalization business' (El Difraoui *et al.*, 2015; De Graaf & Weggemans, 2018) further legitimizes questions of

whether tax-payers' money is being spent effectively. CVE evaluation, by sorting out what is more effective in a given context and how, can help us to develop the academic field, as there is currently not even consensus on a conceptual and theoretical framework guiding the implementation of validated prevention strategies. CVE evaluation can thus improve practice as well avert malpractice. Finally, CVE evaluation can stimulate human rights and compliance with obligations of non-discrimination by shedding light on potential counter-productive effects of targeting specific audiences.

We are in need of an evaluation method that can accommodate the complexity and contextuality of CVE, while providing an authoritative basis for new evaluations and at the same time exposing incorrect or too normative assumptions in CVE policy. The search for methods is made more difficult by the currently limited availability of CVE evaluations, which are essential for theory building. The method sought should thus be able to overcome this constraint. Preferably, it should also support the building of further conceptual understanding, by establishing the 'how' of the success or failure of specific measures.

Realistic evaluation and realist review claim to be designed for just such challenging and complex conditions. Realistic evaluation was developed for the evaluation of social programmes, understood as social systems in which there is an "interplay between individual and institution, agency and structure and micro and macro social processes" (Pawson & Tilley, 1997: 63). Seeking to go beyond a simplistic quantitative-qualitative dichotomy, Pawson and Tilley (*ibid.*) stressed the mechanisms (M) that induce effects, and they introduced the notion of context (C). Their approach stems from the realist proposition that the "relationship between causal mechanisms and their effects [is] not fixed, but contingent" (Pawson & Tilley, 1997: 69). Realistic evaluation aims to identify the combination of mechanisms and context that leads to particular outcome patterns, also known as context-mechanism-outcome pattern configurations or 'C-M-Os'.

These indicate how programmes activate mechanisms, among whom and in what circumstances, to bring about changes in behaviours or events. In short, realistic evaluation revolves around the question 'what works, for whom, how and in what circumstances' (Pawson & Tilley, 1997; Gielen, 2017a: 3). Although there is an upward trend in the use of realistic evaluation, bona fide realistic evaluation of social programmes remains scarce. Realistic evaluation methods, though developed in part as a critique of randomized

controlled trials (RCTs), have been used mainly in the medical domain as an alternative to RCT-based evaluations (Redfern *et al.*, 2003; Tolson *et al.*, 2007; Greenhalgh, 2009; Rycroft-Malone, 2010).

Though the methodological rules of realistic evaluation are not set in stone and still emerging, a few basic principles tend to be followed (Pawson & Tilley, 1997). Realistic evaluation starts with theory on how mechanisms relate to certain contexts and combined produce certain outcomes. Following the positivist wheel of science, a testable hypothesis is deduced from theory and derived and refined inductively from data (Wallace, 1971). The next step is data collection. Pawson and Tilley (1997) do not prefer any specific form of data collection; rather they advocate multi-method data collection. Data is to be analysed by drawing out relevant contexts and mechanisms that contribute to outcome patterns. This is the third step. What ultimately follows is 'programme specification'. According to Pawson and Tilley (1997: 57), "Programmes work (have successful 'outcomes') only in so far as they introduce the appropriate ideas and opportunities ('mechanisms') to groups in the appropriate social and cultural conditions ('contexts')."

The realist approach was applied to CVE by Veldhuis (2012), who used it to identify relevant contextual factors and mechanisms for rehabilitating violent extremists. Her study is important as it highlights the potential of the realistic approach. It also illustrates that a lot more is necessary to make the realist approach applicable to evaluating specific CVE measures. She applied the approach not to *evaluate*, but to *design* rehabilitation programmes. More precisely, she argued that the end result of her research would be a first step on that road. Rather than as a programme theory for an extremist rehabilitation programme, her work should be considered an exercise to illustrate to policymakers that they should develop a theory of change before they implement a programme (*ibid.*: 7). Her research yielded a rudimentary, tentative theory of change based on the first step of realistic evaluation: formulating a set of hypotheses on what might work, for whom, how and in what circumstances. In the absence of research on rehabilitation programmes for violent extremists, she mainly drew on the literature on 'general' rehabilitation programmes (*ibid.*). She observed that other forms of knowledge can serve as inputs for realist theory development, but claimed that systematic reviews and meta-analyses of existing evaluations are most preferable (*ibid.*: 6). She also suggested alternatives for when such evidence is lacking: single studies, but only if they adhere to the highest possible research standards and expert opinion – though she argued that this should

always be integrated with the best evidence from more systematic research. Veldhuis' proposal – that theories of change should be developed only based on systematic reviews or single evaluations that meet the highest research standards – is the exact opposite of Pawson and Tilley's proposal in their realistic evaluation method. According to them, if we truly want to understand how contexts and mechanisms (potentially) produce outcome patterns, we cannot limit ourselves to only systematic reviews that provide a simple answer in the form of a mean effect size to the too simplistic question of 'what works'. Rather, in order to really understand what works, for whom, how and in what circumstances, we must not exclude any research method. They claimed that 'folk theory', in the form of, for instance, expert opinion, policy documents and participant observation, can provide valuable insight into relevant C-M-Os (Pawson & Tilley, 1997).

Yet, unfortunately, while Pawson and Tilley claimed their evaluation method to be suitable for situations in which prior evidence is lacking, they offered little insight as to how, for example, expert opinion can be used to develop a theory that hypothesizes on what works, for whom, how and in what circumstances, which is the starting point of realistic evaluation. The only thing they offered on that point is this very brief description from a case they evaluated:

[W]hilst the case for adult education has been made many times, available theory on its rehabilitative potential was negligible. Research thus began with a period of theory development derived from qualitative investigation, in which educators were interviewed to elicit realistic theories on program mechanisms ('what it was about the programme which might generate change in a prisoner') and program contexts ('with which sorts of inmates in which conditions the initiative might be successful') (*ibid.*: 88).

Using 'folk theories' is not uncommon in the social sciences. Related notions referred to by other scholars are background theories and appreciative systems, which are central in *reflection-in-action* (Schön & Rein, 1994). Webber (1990) discussed a variety of sources of knowledge (e.g., formal scientific and media-based), synthesized by policy actors into policy theories, and others have stressed that this also includes forms of informal, practice-based knowledge, such as *mētis* (Scott, 1998) and *local knowledge* (Yanow, 1996, 2006). Yet, Pawson and Tilley provide no (heuristic) guidelines on how to apply folk theories when conducting realistic evaluation.

In sum, the realistic evaluation method is a promising heuristic evaluation model to apply to CVE, because it can take into account the contextuality and (normative) assumptions that underlie this wicked and contested problem. However, realistic evaluation has never before been applied to CVE. While realistic evaluation offers guidelines on how to use theory to deductively develop a theoretical model in situations where theories and evidence are abundant, it provides no detailed guidelines on how to develop a theoretical model that can be tested in situations where theories and evidence are lacking. This aspect of the realistic evaluation method requires further development.

Realist review

As an alternative to systematic review methods that cannot deal with the challenges of evaluating complex programmes and interventions, Pawson (2002b, 2006) developed the method of realist review, sometimes referred to as ‘realistic review’ or ‘realist synthesis’. Realist review relates to realistic evaluation in the same way that meta-analysis relates to randomized controlled trials in traditional evidence-based medicine and policymaking. Realist review adheres to the same realist principles as realistic evaluation. Like realistic evaluation, the overarching goal is to create a middle-range theory of for whom, how and in what circumstances programmes work. The theory can then be used to formulate policy recommendations, either to hone implementation or to shape new interventions (Pawson, 2006: 74). Realist review differs from realistic evaluation in the research focus. While realistic evaluation is meant for evaluation of a specific programme or intervention, realist review is a method to review and synthesize *existing* evaluation studies. Realist review is helpful where existing studies of complex interventions are heterogeneous or limited. Realist review does not provide a ‘yes or no’ answer or ‘mean effect size’ in response to the typical evidence-based policy question of ‘what works’. Instead, it is explanatory, again revolving around the question ‘what works, for whom, how and in what circumstances’.

Pawson (2006) explained that realist review follows the six steps of the Cochrane reviews. Cochrane is a global database of systematic meta-analyses of interventions in the healthcare sector aimed at improving evidence-based healthcare decision-making. Pawson (2006: 41–42) summarized the six steps of the Cochrane review process as follows: (i) clarifying the scope of the review; (ii) searching for primary studies; (iii) appraising quality; (iv) extracting the data; (v) synthesizing the data and (vi) disseminating the

findings. Realist review differs from the traditional Cochrane review in several ways. For instance, it does not focus on ‘what works’ or on translating outcomes into mean effect sizes, but explicitly takes contextual factors and mechanisms into account. Closely related, a realist review cannot be standardized or reproduced in the same sense as a Cochrane review. Instead, the review is explorative and iterative in nature (*ibid.*: 91–94). More detailed guidance on how to conduct a realist review can be found in summary articles (Pawson *et al.*, 2005; Wong *et al.*, 2012) and empirical studies drawing on realist review (O’Campo, 2009; Wong *et al.*, 2010). Most realist reviews have been done in the field of healthcare. Researchers conducting such reviews can often draw on a large body of literature for hypothesizing what might work and how, but also when searching for primary studies for review (Gielen, 2017a: 4).

With the exception of Veldhuis’ work, realist review has never before been applied to CVE. Researchers using the approach in this field face the lack of an abundant body of literature to draw on for meta-analysis. Pawson (2006), however, has claimed that realist review is also appropriate for areas of intervention and programmes in which evaluation is lacking. Having said that, Pawson (2005) also indicated that the realist review method would require adaption and innovation in order to address the widening range of complex programmes (Pawson, 2006: 93–96 in Pawson & Manzano-Santaella, 2012).

This need for further development of the realist review method to accommodate increasingly complex programmes was also pointed out by Betts (2013). Her realist review of aid effectiveness and governance for developing country reform concluded that broad and diverse programmes and interventions require a more systematic approach for data extraction (*ibid.*). So in theory, realist review can provide a suitable method for synthesizing the limited CVE evaluation literature, contributing in turn to theory building. This would provide a more solid basis for new CVE evaluation. It is, however, likely that the review method would have to be adapted to be applied to CVE.

In sum, with Veldhuis’ work being an important exception, neither realistic evaluation nor realist review have been previously applied in relation to CVE. More generally, too, these techniques are still under development. Pawson and Manzano-Santaella (2012) observed that few evaluations and reviews labelled as realist actually abide by the realist principles. This is also true of the study by Veldhuis. While not stated as a realistic evaluation, she

proposed research methods for theory development that are substantially aligned with realist principles. So, while on paper realistic evaluation and review seem suitable for tackling the evaluation of CVE, in practice the method might need additional development. The further development of this evaluation method in order to make it suitable for the evaluation of CVE lies at heart of my research, as presented in this dissertation.

1.5 Research focus and approach

To reiterate, CVE is a relatively new and very complex phenomenon. The term itself has become somewhat of a catch phrase for a broad array of interventions targeting different populations at different stages in the radicalization process. Most common evaluation methods have proven unsuitable for dealing with the complexity of CVE. The result is a serious lack of evaluations in the field. Academia thus has little theoretical basis for adequately advising the policy and practitioner community on the best course of action. Considering the social impact of violent extremism and the economic consequences involved, the CVE ‘industry’ has become the brunt of criticism, and both CVE interventions and practitioners are now being subjected to increased scrutiny.

Greater focus on the evaluation of CVE is therefore called for. Realistic evaluation seems to be a good means of dealing with the complexity of CVE programmes. However, applying realistic evaluation to CVE is unprecedented. Will it indeed prove to be a suitable method, or will adaption be required? Also, can it help us gain a better understanding of CVE, ultimately contributing to the building of a theoretical foundation upon which the research community can base its policy recommendations? This leads to the central focus of this dissertation:

How can realist review and realistic evaluation be employed and developed in the field of countering violent extremism (CVE), and what can realist review and realistic evaluation contribute to the development of CVE policy and to deepening our understanding of violent extremism?

This focus can be divided into the following five questions:

1. What is countering violent extremism (CVE) (chapter 2)?
2. How can realist review be employed and developed within the social sciences and, in particular, in the domain of CVE (chapters 2, 3 and 6)?

3. How can realistic evaluation be employed and developed within the social sciences and, in particular, the domain of CVE (chapters 4, 5 and 6)?
4. How and what can realistic evaluation and realist review contribute to development of more informed CVE policy and programmes (chapters 2 and 6)?
5. How and what can realistic evaluation and realist review of CVE measures contribute to deepening our understanding of the phenomenon of violent extremism (chapters 4, 5 and 6)?

This dissertation brings together a number of articles either published or under consideration by international peer-reviewed journals. Chapter 2 is based on the published article *Countering Violent Extremism: A Realist Review for Assessing What Works, for Whom, in What Circumstances and How?* (Gielen, 2017a). It presents a realist review of existing CVE evaluations, which provides us with a better understanding of what CVE is and dos and don'ts for CVE policymakers and practitioners. It thus helps us to answer the first, second and fourth research questions.

Chapter 3, *Lessons from a Realist Review of Countering Violent Extremism: Development of Refined Heuristic Guidelines for Complex Social Programmes*, offers a secondary analysis of the realist review presented in chapter 2, concentrating on the methodological and heuristic aspects of conducting a realist review. In contrast to chapter 2, chapter 3 does not focus on the *outcome* of a realist review of CVE measures, but rather on the steps toward it and the challenges that arise when applying the method of realist review to CVE. Heuristic guidelines for realist review of complex social programmes (such as CVE) are presented, contributing to the answering of the second research question.

Chapter 4 is based on the published article *Supporting Families of Foreign Fighters: A Realistic Approach for Measuring Effectiveness* (Gielen, 2015a). It zooms in on one specific measure: family support. Family support is a relatively new measure under the CVE prevention umbrella. The chapter elaborates how folk theory can be used to develop hypotheses on context-mechanism-outcome pattern configurations (C-M-Os) as part of the first step of realistic evaluation. The chapter illustrates how family support programmes might activate mechanisms, among whom and in what circumstances to generate change. The development of C-M-Os based on folk theories is an important step in applying and developing the realistic evaluation method for CVE (research question 3). The developed hypotheses

help us to gain a better understanding of the role family plays in radicalization and deradicalization processes (question 5).

Chapter 5 is based on the published article *Exit Programmes for Female Jihadists: A Proposal for Conducting a Realistic Evaluation of the Dutch Approach* (2018a). In contrast to chapter 4, it goes beyond the first step of realistic evaluation of developing rudimentary C-M-O configurations, to illustrate how all four steps of realistic evaluation can be applied (question 5). Exit programmes for Dutch female jihadists are taken as a case study. The chapter examines how the proposal for conducting a realistic evaluation can contribute to more informed policymaking on exit programmes (question 4) and to theory development on female jihadists (question 5).

The focus of the above chapters is on the application of realistic evaluation and review in ex post situations. Chapter 6, *Evidence-Based Informed Policy Design for Support Groups for Families of Foreign Fighters: Ex Ante Application of Realistic Evaluation and Review*, addresses the question of whether and how realist review and realistic evaluation can also be applied ex ante, in order to develop a more evidence-based and informed CVE policy theory and design. This chapter provides a description of a new CVE intervention: support groups for families of foreign fighters. It illustrates how both realist review and realistic evaluation can be applied (question 2 and 3) to support groups for families of foreign fighters, leading to a more evidence-based policy theory. The chapter goes further than previous chapters in that it not only develops an evidence-based policy theory, but actually tests that theory in practice by interviewing and organizing focus groups for families of foreign fighters, several professional and community organizations and policymakers (question 3). This chapter demonstrates that such an ex ante approach can contribute to legitimacy, feasibility and effectiveness of CVE policy (question 4).

Chapter 7 returns to the central focus of the dissertation. Drawing on the findings of the previous chapters it provides answers to the larger questions of how realistic evaluation can be employed and developed in the field of countering violent extremism (CVE), what realistic evaluation and realist review can contribute to the development of CVE and how realistic evaluation and realist review can contribute to deepening our understanding of violent extremism. It also presents several recommendations for the improvement of CVE policymaking and evaluation.

1.6 My role as a researcher

Throughout this introduction I have noted that some of my insights and case studies arose from my work as an independent researcher and consultant (A.G. Advies), which I have been doing for more than a decade. After graduating as a political scientist in 2007, I embarked on an evaluation journey in which I attempted to assess programmes and interventions aimed at preventing radicalization and violent extremism. Time and time again I found myself confronted with the limitations of 'regular' evaluation methods, such as experimental evaluation (RCTs), process evaluation and interpretive evaluation.

From 2012, I worked as an independent consultant for the Radicalisation Awareness Network (RAN) of the European Commission, for which I was involved in the RAN working groups on internet and social media and internal and external dimensions (foreign fighters). The latter working group was dedicated to the European foreign fighters travelling to Syria. We organized meetings twice a year with policymakers, intelligence and security professionals, and practitioners and community organizations that were dealing with the foreign fighter phenomenon. At these meetings participants shared their views on both the phenomenon of violent extremism and particularly foreign fighters and also shared the policies and interventions they had designed and implemented to prevent and counter violent extremism. As an evidence-based policy to prevent and counter the foreign fighter phenomenon was lacking, the policies and interventions were mainly based on the assumptions of policymakers, practitioners and community members or their view of the problem and how the problem could potentially be tackled. These policy theories, were rooted in years of experience and knowledge exchange with policymakers and practitioners across Europe.

It was at one of these meetings, in April 2013 to be precise, that intelligence and security professionals observed that an alarming number of people who were previously part of criminal gangs had travelled to Syria, making use of their criminal experience and network in Syria (e.g., their ability to access and use weapons). While this *crime-terror nexus* had been studied in academia (e.g., Bibes, 2001; Shelley & Picarelli; Makarenko, 2004), it was not until three years later, in 2016, that it was addressed in relation to European foreign fighters (Basra *et al.*, 2016). So while the academic world was and still is trying to grasp the foreign fighter phenomenon in terms of who foreign

fighters are and what drivers, root causes and trigger events caused them to travel to a combat zone, policymakers and practitioners had to make sense of the situation, while working to prevent and counter it at the same time. This led to the sharing of interesting insights and practices.

In September 2013, another RAN meeting was organized which revolved around the question of what practitioners and communities were doing to prevent and counter the foreign fighter phenomenon. It was at this meeting that the situation of families of foreign fighters was addressed. Practitioners and communities from different European countries had noticed that the family members of foreign fighters were heavily impacted. In Germany, where practitioners had a history of dealing with families of right-wing extremists, the first support programmes for families of foreign fighters were being offered. In Antwerp, one policy officer decided to literally go out knocking on the doors of the affected families to ask them how the municipality could help. In Denmark the city of Aarhus decided to organize support groups for families so that they could share their experiences. All of the practices described offered insights into relevant contextual conditions, the mechanisms behind offering family support and its potential and real outcomes, which were summarized in meeting reports. I was responsible for taking those minutes.⁷ I then sent those minutes to all of the participants at the meeting to confirm that my report correctly reflected and summarized the most important analyses and conclusions of the meeting. The reports of the RAN foreign fighter working group have since led to several publications authored or coauthored by me on supporting families of foreign fighters (RAN, 2013; Gielen, 2014).

An unintended consequence of my publications on family support was that one of my articles (Gielen, 2014) became the first hit on Google searches for 'support for families of foreign fighters' (in Dutch 'familieondersteuning voor Syriëgangers'), as at the time no publications were available on the families of foreign fighters. It was not until 2016 that the first scientific publication was published on this topic (Maher & Neumann, 2016). Several months after the 2014 publication I received a phone call from a social worker. He had been approached by the father of a foreign fighter and wanted to help him, but did not know how. He found me by conducting a Google search on how to support families of foreign fighters. Based on my experience within RAN

⁷ These minutes are available upon request to the RAN host organization, RadarAdvies in Amsterdam, or the European Commission Directorate-General for Migration and Home Affairs.

and the foreign fighter working group, I started by asking him some very basic questions such as when the son had left and whether the police had been notified. The social worker eventually decided he felt out of his depth and asked if he could refer the father directly to me. Within two minutes I received a phone call from the father, who shared the story of his son who had travelled to Syria. The father, it turned out, had been sent home from the police station when he wanted to file a report of his son as a missing person. This shocked me, because if the police refused to register foreign fighters, how could the relevant organizations know who had gone – let alone returned (with combat experience). As an advisor to the Dutch central government and municipalities, I had built up a relevant network of people who I contacted to raise the issue. This resulted in a community policy officer coming to the father's house so that he could – discretely – report his son as a missing person.

I anticipated that my role would end there. The father had a specific problem, which I helped him with, and that was the end of it. However, as there was no organization in the Netherlands at that time that offered family support, the father had no one else to turn to but me. As he felt I was knowledgeable on the issue, he not only phoned me to ask additional support questions, but also referred other families of foreign fighters to me. I thus unintentionally became a support provider to families of foreign fighters. I initially did this on a voluntary basis next to my PhD studies and work as an independent researcher and consultant to RAN, ministries and municipalities. However, this soon became a full time job and I was no longer able to help new families. I approached the ministry and municipalities to inform them of the parents' strong need for family support and the fact that there was no organization in the Netherlands that could provide the support. In addition, I could no longer help new families on a voluntary basis. From that point onwards, I continued to support the families that were already part of my 'voluntary caseload'. But new families that needed support would have to be financially covered by their municipality. Helping families also resulted in municipalities asking me to provide interventions for young females who had tried to leave for Syria. Providing support to families of foreign fighters and to individuals who had attempted to travel to Syria thus became part of my independent consultancy work. As part of this work I took notes and wrote reflections on my meetings with these families and individuals in my professional notebooks. The notebooks were archived in accordance with the code of conduct as developed by the Association of Universities in the Netherlands (VSNU Gedragscode voor Gebruik Persoonsgegevens

in Wetenschappelijk Onderzoek). As a rule to municipalities, I worked with a strict privacy protocol: the information families shared with me and the notes I took were strictly confidential, unless families gave their explicit permission for me to share their answers to specific questions or experiences. The municipalities in question were aware of my assistance to these families. Therefore, the actual existence of the cases, can be verified with the municipalities.

While I initially saw my work as a family support and intervention provider completely separate from my PhD, the work brought me many insights that – at least at the time – could not be found in the scientific literature on CVE, on family support and on exit programmes. In essence, my work provided insights on the target audience of CVE programmes, information on the types of programmes and interventions that were being implemented by practitioners and policymakers and the lessons they and I had learnt during implementation. It thus helped me to better understand the underlying assumptions (mechanisms) of CVE interventions and relevant contextual conditions that produced certain outcomes. In short, my work uncovered folk theories and valuable data for the first step of realistic evaluation: developing theories (hypotheses) on what *might* work, for whom, how and in what circumstances. In other words, I was suddenly provided with a key to my evaluation puzzle. The insights I collected at conferences and meetings with policymakers and practitioners, as well as my voluntary and consultancy support to families of foreign fighters and their radicalized children, and had reported on as a *reflective practitioner* (Schön, 1983) proved invaluable as inputs for theory development, which was the first step of the realistic evaluation method. While the insights from my voluntary and consultancy work with relatives of foreign fighters and radicalized individuals are included in this dissertation, it is important to note that my analysis does not critically rely on them. They are included in chapter 4 (on families of foreign fighters), chapter 5 (on exit programmes for females) and chapter 6 (support groups for families of foreign fighters). These chapters describe how the data gathered as a reflective practitioner may be used in the first step of realistic evaluation as part of constructing a tentative theory to be tested in later steps. In other words, in these three chapters, the data was never used to draw conclusions about the effectiveness of specific programmes or interventions. Rather, the data was used only as part of the first step of realistic evaluation: developing hypotheses on what *might* work, for whom, how and in what circumstances. In chapter 4 the folk theories helped me to build C-M-O configurations for family support which

can be used by other scientists for the actual evaluation of family support programmes. In chapter 5, one (anonymized) case study of a radicalized female was used to illustrate how case studies can be analysed in a realistic fashion so that more academics can apply the method, ultimately leading to greater insight into the effectiveness of programmes aiming to disengage and deradicalize. In chapter 6 multi-method data collection was applied, of which interviews and focus groups with families of foreign fighters was only one part, to test the hypotheses formulated via realist review. This research had complete autonomy in formulating the research questions, in choosing the research methodology and in selection of the research participants.

To conclude, my work as a reflective practitioner helped me to develop the realistic evaluation method, in order to make it applicable to CVE evaluation, which was precisely the aim of my PhD research.