The Data Collection Challenge:
Experiences Studying Lone-actor Terrorism

Paul Gill
The views in this publication are those of the author. They do not necessarily reflect the views of the RESOLVE Network, its partners, the U.S. Institute of Peace, or any U.S. government agency.
ABSTRACT

In 2011, the United States (U.S.) Department of Homeland Security (DHS) awarded our research team a one-year project to build a database on the antecedent behaviors of lone-actor terrorists from scratch. This chapter outlines the variety of challenges we faced from the outset of the project, how we tried to address them, and how I would do things differently if we started all over again. In particular, the chapter focuses on (1) the trade-off between definitional ideals and pragmatic constraints, (2) the limitations and opportunities afforded by a reliance on open-sources (and what I’ve since learned about closed-sources), (3) deciding what to collect data on and how to code what we collected, and (4) how to structure the data collection team and implement internal quality control mechanisms. The chapter then demonstrates how these decisions affect the quality and prevalence of certain variables by comparing our results to analogous studies. The chapter concludes by outlining important considerations for pursuing and understanding similar projects.

INTRODUCTION

From January 2010 to late 2012, I was a postdoctoral research fellow at the International Centre for the Study of Terrorism (ICST) at Pennsylvania State University under the Directorship of Professor John Horgan. My initial task was to lead the day-to-day research of a project entitled From Bomb to Bomb-maker, sponsored by the Office for Naval Research. Prior to starting the job, John Horgan shared the project proposal with me. A large component of the proposal promised systematic data collection on Provisional Irish Republican Army (PIRA) bombings, core PIRA members, and their network links. It seemed like a very ambitious endeavor and the ever-present imposter syndrome was particularly salient before I even began employment there.

What Prof. Horgan did not mention in early communications was that he had developed a competitive internship program at the Centre. Enthused, energetic, and diligent undergraduate students applied en masse to work ten hours a week on whatever the Centre’s research required. Some of those undergrads applied to obtain course credit, others purely for the experience. With this type of research support, suddenly things looked more doable. In that first year, dozens of students helped us collect data on over 5,000 PIRA bombings, 1,200 PIRA members, and the thousands of network connections between them.

During data collection, we learned many lessons, including four detailed in this chapter:

- Data availability, research feasibility, and definitional choices,
- The strengths and limitations of purely open-source data collection methods,
- Data collection and developing and testing a coding schema, and
- Quality assurance and structuring coding teams.
While it was clear that open-source efforts have a lot of promise, their utility to the research field depended on the types of questions asked of the data and the degree of organization, quality control, and manpower in the data collection process.

In March 2011, during my tenure at ICST, DHS awarded our research team a one-year project to build a database on the antecedent behaviors (those occurring prior to an individual carried out an attack) of lone-actor terrorists from scratch. The project aimed to explore lone-actor terrorism from both individual and event-driven perspectives through several rigorous data-driven analyses. This included synthesizing research on analogous cases of lone actors, creating a typology of lone actors, and cataloging their pre-attack behaviors.

We identified multiple benefits of this research for informing the development of operational considerations. First, it would provide the first empirically grounded, methodologically rigorous, multidisciplinary, and multi-level analysis of lone-actor terrorism from open source data. Second, it could provide an empirical basis—derived from analogous cases and detailed case studies—for identifying recurring pre-attack behaviors. Finally, it could inform security planning through providing an empirically informed understanding of the behaviors associated with and leading up to lone-actor terrorist attacks. It is hoped that reflections on the project as outlined in this chapter will not only provide additional transparency to our research process (as we hope to make the data fully available soon), but also guidance for future researchers setting out on similar data collection endeavors.

PROJECT BACKGROUND

The mission of ICST was to engage in and promote the scientific study of terrorism and political violence. In doing so, ICST sought to create multidisciplinary, cross-national research teams, drawing strongly but not exclusively on the social and behavioral sciences to respond to needs and opportunities in the broad research areas of terrorism and counterterrorism. The Center was committed to engaging in policy relevant, yet independent and non-partisan research. The overarching mission of the ICST is to provide “actionable knowledge,” defined as a conceptual basis to policy-relevant and operational counterterrorism activity.

A by-product of this mission was that we came into contact with many practitioners who would either visit us on campus or who we would see on regular trips to conferences and workshops in Washington, DC. The phenomenon then popularly known as lone wolf terrorism was a recurrent theme in these conversations. This interest in lone wolf terrorism was partially event driven. During those initial 15 months at ICST, several incidents occurred in the U.S. that appeared to fit the moniker. They included the deliberate crashing of a plane into an IRS building in Texas, the shooting at the Pentagon Station, the shooting of police officers in West Memphis, the Times Square bombing attempt, the hostage taking at the Discovery Channel’s headquarters, the Northern Virginia military shootings, and the attempted bombing of
a Martin Luther King Jr. memorial march. Additionally, several FBI investigations led to successful and well-publicized disruptions of so-called would-be “lone wolfs” (although in many cases the offenders in question thought they were part of an actual terrorist cell).

Academic perspectives on the topic at the time remained under-researched. What existed was largely weak on methods, theory, and empirics, and, therefore, largely ill-suited for the practitioner community. This, coupled with the sense that lone wolf incidents were on the rise, and our own self-efficacy in systematic open-source data collection, pushed us to design a research project to address the disconnect between research, policy, and practice through a project examining lone actor behaviors.

Initially, my cynical, and uninformed Irish self felt the “lone wolf” problem was largely an American one. I presumed easy access to firearms, comparatively poorly regulated improvised explosive device (IED) component markets and limited but costly access to mental health care contributed to an environment prone to “lone wolf” incidents unique to the United States. While this may be true, it does not explain why similar phenomena occur in other parts of the world. On a more personal level, I also wanted to further test a basic premise underlying my PhD research: the influence of group and social dynamics on extremist behavior. Lone actors therefore seemed to be the perfect opportunity to test my framework of understanding.

While writing our proposal for DHS in early 2011, a cursory search showed many similar European incidents that had largely eluded the US headlines, including attempted assassinations of the Danish cartoonist Kurt Westergaard and British Member of Parliament Stephen Timms, a bombing of a Danish newspaper office in Copenhagen, a coordinated car and suicide bombing in Stockholm, and a shooting at a Frankfurt Airport targeting U.S. Air Force personnel among others. It was immediately clear that lone attacks were not exclusive to the U.S., and that there was more to the phenomenon than initial lazy assumptions suggested. We pitched our project to DHS in March 2011, responded to comments in May 2011 and started work in August 2011. One month prior to starting, the right-wing extremist Anders Breivik killed over 70 people in separate bombing and shooting attacks on a single day in Norway.

LESSON 1: DEFINITIONAL IDEALS & PRAGMATIC CONSTRAINTS

While designing the project proposal, one of the first questions we faced was who are we studying. Getting a discrete definition down on paper would set the parameters for everything that followed, make all the subsequent steps in our research process (as detailed below) more manageable, and provide a

3 Gill, 2015; Gill, Lone-actor Terrorists.
foundation for later data inclusion and exclusion-related questions. It was important to get our initial definition right, and we dedicated a substantial amount of time to doing so before data collection began.

Choosing Terms

The project proposal eschewed the frequently used term lone wolf terrorism for multiple reasons. First, in the animal kingdom, wolves become “lone for short periods as they move from one pack to another for mating reasons (rather than simply going it alone forever).4 Second, wolves are usually associated with cunning, intellect, and danger—traits that should not be applied to often incompetent attackers whose actions largely depend on luck. Third, the phrase has a connotation of “coolness” that may motivate some individuals to become a “lone wolf” attacker themselves. Finally, as my colleague Bart Schuurman once said, the term is also offensive to wolves.

Ultimately, we circled around the term lone-actor terrorism and stuck with it. To be clear, there was no shortage of alternatives. Frequently utilized terms include “freelancers”, “lone operator terrorist”, “solo terrorists”, “solo actor terrorists”, “loners”, “stray dogs”, “lone offenders”, “menacing loners”, “leaderless resistance”, “leaderless terrorism”, “individual terrorism”, “single actor terrorism”, and “self-activated terrorism.”5 Most important when selecting a term, however, is that it fit the nature of the phenomena to be studied to the greatest degree possible and is coupled with a clearly circumscribed and articulated definition.

Deciding on Definitions

Initially, our proposal neglected to define what we understood lone-actor terrorists to be. In the weeks between hearing of the proposal’s success and the project’s start, we spent a lot of time juggling definitional ideals and pragmatic constraints. Some studies on lone actors excluded individuals with any

---

4 Gill, 2015.
connections to a broader network, while others allowed for their inclusion. Some studies used definitions allowing only one person’s involvement in a terrorist plot, while others allowed for the inclusion of “isolated dyads” or even more accomplices in the plot. Still others went as far to suggest a “lone wolf pack” to include many more individuals that operate within a co-offending network, but lacking in formal connections to a formalized terrorist command-and-control structure. Some definitions restrict the observation pool to those inspired by specific ideologies, while others allow the inclusion of those who carried out attacks because of personal or criminal motivations. Some interpretations include offenders whose political motives were not immediately obvious, while others highlight the importance of who was targeted. Some, like Feldman, stipulate that non-military targets must be the target of the attack, while others, like Kaplan, allow for the targeting of government agents or buildings. Some allow for lone actors to be directed in some way by a wider group of individuals or a movement, while others explicitly reject this component. Finally, some definitions exclude individuals with a history of mental illness.

Ultimately, some of our definitional decisions were made for purely research purposes. We decided individuals needed to be ideologically inspired, but we chose not to restrict data collection to any one ideology. For the purposes of the project, we borrowed from official United Kingdom (UK) parlance with terrorism defined as the use or threat of action designed to influence the government, intimidate the public or a section of the public, or advance a political, religious or ideological cause. Terrorism can involve violence against a person, damage to property, endangering a person’s life (other than that of the one committing the action), creating a serious risk to the health or safety of the public or a section of the public. In line with the UK’s definition, we did not want to exclude cases based on the target of the attack. Adopting a definition that mandated an ideologically inspired action but did not restrict based on the target of that action, we determined, would allow for the possibility of comparative analysis across ideologies and across targeting strategies.

8 Spaaij, Understanding Lone Wolf Terrorism.
10 Hewitt, Understanding Terrorism in America.
12 Bakker and de Graaf, “Preventing Lone Wolf Terrorism, 8.
15 Feldman, “Comparative Lone Wolf Terrorism, 270–86.
17 Feldman, “Comparative Lone Wolf Terrorism, 270–86.
18 Burton and Stewart, “The ‘Lone Wolf’ Disconnect.”
19 For the official UK definition, see: https://www.cps.gov.uk/terrorism.
Pragmatic Constraints

Other definitional decisions we made were for purely pragmatic purposes. As additional project outputs, we promised a descriptive analysis and a quantitative analysis from which to derive a lone-actor terrorist typology. This necessitated a larger sample size than what a more restrictive definition would allow. While lone actors (e.g. those operating completely independently of a group yet claim the attack on behalf of a group/movement or, alternatively, those that may have radicalized within a wider group but subsequently left and carried out terrorist acts outside of a formal command and control structure) would form the core of our data collection, we needed to extend our definition a bit to increase our sample size. We, therefore, decided to also include “lone dyads” (e.g. pairs of individuals who operate independently of a formal group, conceiving, developing, and carrying out terrorist related activities without the input from a wider network) and “solo terrorists” (e.g. individuals who are specifically trained and equipped by a terrorist group, but who attempt to carry out the actual attack alone) to broaden the sample size.

Conceptually, these two categories were similar to and maintained the integrity of more traditional definitions of lone-actor terrorists both in terms of undertaking an attack on their own (in the case of “solo terrorists”) and in terms of having been inspired to carry out an attack outside of the confounds of an actual group (in the case of the “lone dyads”). Put simply, both “lone dyads” and “solo terrorists” each operated in a manner one step removed, independent of, and/or divested from larger terrorist operations and group command structures. Based on our PIRA data collection experience, we stopped short of including “triads” (groups of three individuals involved in a terrorist attack absent commands or input from a broader terrorist group) and larger co-offending networks (or “lone wolf packs”) because of serious concerns about the availability of data given the involvement of multiple plotters in an attack, an issue discussed further later in this chapter.

We also made the pragmatic decision to specify that the individual needed to be lone at the time of the attack rather than throughout the radicalization process. Determining whether an attacker acted alone during the attack itself is simple. It is a discrete moment, a single day, and therefore much clearer and less resource intensive to identify and study. A consistent problem arose, however, in cases where a single individual was arrested and convicted for plotting an attack, giving the appearance of relevancy, but in which it later became clear that other individuals, notably confidential informants, played a key role in shepherding the plotting of an attack. In these cases, the sole arrested individual may have sincerely felt as though they were a part of a group in plotting the attack, and therefore, given our own definitions, would not fall within the category of a “lone actor.”

Data Availability

For data availability issues, we additionally restricted our data geographically to those who planned their attack within the “West” (e.g. Europe, North America, Australia, New Zealand); there simply was not enough material and language proficiency available on non-Western cases. Fortunately, subsequent collaborative projects have demonstrated some of the ecological and cultural validity of our findings on
Palestinian lone actors. The field could benefit from further collaborations to build out this study in additional contexts.

Knowing that each of these decisions had the potential to sway results in any given direction, we also decided to code each of these important criteria to allow for alterations in our definition. Doing so granted us flexibility in later adapting the research using a new definition (e.g. the author’s later decision to exclude solo terrorists and lone dyads from the definitional criteria) or in cross-checking validity by comparing results among different definitional permutations (which is now a relatively common request).

Prior to data collection, and after narrowing our definitional criteria, we also examined previous academic literature on lone-actor terrorism to construct our research or “actor” dictionary – a list of offenders fitting our previously established criteria. Beyond previous literature, we sourced additional offender names through tailored search strings and applied to LexisNexis. Initially, we were restricted to English language resources; however, in later projects we were able to collaborate with Dutch, German, Danish, and Polish colleagues with the requisite language skills. More individuals were also identified through the Global Terrorism Database (GTD) developed by the National Consortium for the Study of Terrorism and Responses to Terrorism (START), as well as through lists of individuals convicted of terrorism-related offenses.

During our attempts to construct an actor dictionary, however, we ran into another data constraint. The data available through LexisNexis—upon which we primarily relied when populating our dictionary and dataset—only contained limited information on terrorist events that occurred prior to 1990. Even the more basic task of simply identifying lone actors that did not dominate the news prior to 1990 is a rather difficult challenge given the relatively limited records from decades prior. As such, we decided to further limit the population of the study to post-1990 events and offenders.

Reflecting on the definitional choices, parameters, and constraints that limited the scope and reach of our study, a few key points are particularly salient. First, regardless of the subject of a particular study and the seeming relevance of the data gathered within it, both researchers and consumers of research should pay close attention to foundational definitional choices and their impact on database construction and the information housed within said database. It is not hard to imagine two parallel studies on the same topic setting forth two very distinct sets of contrasting results based on the inclusion and exclusion criteria established in their initial definitions. We return to similar issues in the concluding section of this chapter. Second, in many cases, the definitions chosen will be dictated by both pragmatic and data-related constraints. Researchers and research consumers should pay close attention to how the availability of data may limit the research query and, therefore, research findings.

LESSON 2: RELYING ON OPEN SOURCES: OPPORTUNITIES & LIMITATIONS

As I noted earlier, my early ICST experiences convinced me of the utility of open source data, which proved even more useful when studying lone actors. Given the particularly low rate of, and therefore heightened attention to, lone-actor terrorist attacks, the volume of granular, behavioral data made available through public media and reporting on individual lone actors tends to be much higher as compared to actors who operate on behalf of a specific group and/or within ongoing campaigns of violence where trials and convictions are a weekly occurrence. For example, we were able to obtain educational data for 65% of the our lone-actor sample. This is, compared to less than 10% of our sample of Provisional Irish Republican Army (PIRA) members, for whom level of education could generally only be inferred from the individual’s occupational status.\(^\text{23}\) The relatively higher level of granularity in data on lone actors made available through the media and open reporting makes open source materials a particularly rich repository of information with which to study and understand individual behaviors and characteristics.

Open-Source Data Opportunities

How does open-source data compare to closed-source data? Does it provide sufficient information on violent extremists? Following our research for DHS, our research team carried out another project using closed-source data in collaboration with the Greater Manchester Police.\(^\text{24}\) Findings from this closed source study replicated broadly similar results as identified in other studies that utilized open source data collection efforts. Based on this, we might conclude that outcomes from open source data, when properly collected and handled, may be more accurate than commonly perceived. However, it is possible that this may only be true in studies on lone actors in which, as noted earlier, openly available data tends to be far richer and more closely covered by the media.

In addition, because open-source data is freely available to the public, open-source collection initiatives also offer the bonus of research process replicability—wherein other researchers are able to replicate the exact methods of data collection and analysis used in another study—rather than just attempt to test or replicate research outcomes. Replicating research processes can be more difficult when using a completely closed-source data collection protocol, given that access to the data necessary to replicate the study is not widely available.

Collating closed-source data, moreover, can—and proved to be—an onerous and time-consuming task, one surprisingly more complex than comparable open-source initiatives. Missing data during closed-source collection was a common problem for many of the same reasons it is in open-source initiatives.

---


Complicating our own closed source data collection efforts was that, in our Manchester study, each UK police force maintained their own internal data and reporting systems in idiosyncratic ways, which slowed the process of merging the data into a single repository. Despite this, however, closed-source data provided more granular insight than open sources in areas like leakage (e.g. instances in which lone actors told third parties about their attack plans), bystander observations, and information on engagement with extremist materials. This level of granularity in closed sources supplied us with the information we needed to more fully examine our questions and behaviors of interest among lone actors than that in open sources alone.

**Open-Source Data Limitations**

Despite its promise, open-source data is not without its challenges. First, open source samples tend to only include information on individuals who planned or conducted incidents reported in the media. It is, therefore, possible that open source data contains important gaps when it comes to incidents that either (a) led to convictions but did not register any national media interest or were only reported in more local sources that were not included in the LexisNexis archives (in the case of our research specifically), or (b) were intercepted or disrupted by security forces without a conviction being made or word of the incident getting out. Second, the level of detail in open-source reports varies significantly across incidents and across publications, limiting data collection to only what is available and can reasonably be collected for each case.

Finally, when using only open-source information, it is often difficult to distinguish between missing data and variables that should be coded as a “no”. Given the nature of newspaper and open-source reporting, it is unrealistic to expect each biographically-oriented story of a lone actor to contain lengthy passages that detail every variable or behavior the offender in question did not exhibit (e.g. the offender was not a substance abuser, a former convict, or recently exposed to new media influences, etc.). Definitive “no” answers were a rarity (less than 5%). This percentage was generally uniform across most variables we included. In our project, we usually coded answers as a definitive “no” in situations where inaccuracies were reported earlier in the news cycle and later rectified about a particular offender.

Our work with the Greater Manchester Police suggests that regularly finding definitive “no” answers may necessitate access to more in-depth closed-source data. Therefore, as a necessary practice, we typically treat each variable in the analysis dichotomously—e.g. the response is either a definitive “yes”, or not enough information to suggest a “yes” and, therefore, a “no”.

It is important for researchers and research consumers to realize and accept the fact that no dataset is perfect. The key is to approach databases skeptically, with an eye towards detail. Researchers building databases, in particular, should work to mitigate potential source limitations, triangulate resources, and again provide the transparency necessary for outside readers to contextualize the data presented and the findings drawn from it.
LESSON 3: DECIDING ON VARIABLES & DEVELOPING A CODING SCHEMA

Overall, deciding the variables to collect data on in each of the individual cases proved our biggest project challenge. Ultimately, we ended up compiling variables from a number of places into a codebook to systematize and guide our interns’ data collection and coding efforts. To do so, first we pulled pertinent socio-demographic questions from other ICST data collection projects. Second, we read everything available on lone-actor terrorism and formulated questions for the codebook based on the hypotheses within these publications to test the validity of some commonly held assertions within the literature.

Third, and perhaps most importantly, we looked to fields outside of traditional terrorism studies. Notably, we read Fein and Vossekuil’s “Assassination in the United States” (although consumed and obsessed over the elegance of their paper may be more precise). Although focused on individuals posing a risk of violence to public figures, the study offered parallels to how we could frame questions to understand lone-actor terrorist behavior. The authors demonstrated how risks could be conceptualized as “the end result of an understandable, and often discernible, process of thinking and action,” and laid the groundwork for defining and measuring every step in the mobilization to violence. With this as our guide, all we needed to do was characterize and map each potential step in said “process of thinking and action.”

Fein and Vossekuil’s paper also provided a gateway into the wider and thriving threat and risk assessment and management literature, which had, until then, largely been siloed from terrorism studies. This introduction was influential to our data collection and later analyses to the extent that my current research group now conducts public figure threat assessment research alongside colleagues at the Fixated Threat Assessment Centre. Influenced by Fein and Vossekuil, we even decided to publish our initial lone-actor paper in the same outlet, The Journal of Forensic Sciences, rather than a traditional terrorism-oriented publication. It is important that terrorism researchers look to, draw insights from, and engage in fields outside of terrorism studies alone. There are plenty of fields that have already constructed frameworks and drawn insights that are highly relevant to the field of terrorism studies—we do not always have to reinvent the proverbial wheel. Likewise, there is plenty of insight that other fields can draw from our work on terrorism studies.

Once the codebook was largely constructed, we piloted our data collection on a small number of lone-actor cases and made notes of behaviors that appeared relevant but were not in our initial codebook. No

26 Ibid., 321.
27 Ibid.
28 The Fixated Threat Assessment Centre (FTAC) is “the first joint NHS/police unit in the United Kingdom. Its purpose is to assess and manage the risks from lone individuals who harass, stalk or threaten public figures.” For more on FTAC, see: https://www.fixatedthreat.com/ftac-welcome.php
The codebook is perfect; having the time to pilot the codebook increases the chances of catching obvious and less-than-obvious issues that may otherwise arise. Testing the codebook also gave us a sense of how long it would take teams to code cases so we could estimate the workload accordingly.

The first case we looked at was Daniel Tovey, a right-wing extremist in the U.K. with apparent plans to ignite a “race war”. Understanding Tovey was complicated and, in many ways, personified our whole experience studying lone-actor terrorism. Evidence gathered by the police suggested that Tovey was planning to target ethnic minorities in what appeared to be a campaign of attacks. Police found an arsenal of weaponry and bomb-making materials in his home prior to an attack occurring. In our review, we identified many factors we expected to see in a lone-actor case. He lived alone, he was not very social, he had not affiliated with any right-wing groups, and there was some serious family trauma in his past. But then came some findings that were a bit unexpected in that they were largely absent from literature on lone actors at the time.

For instance, Tovey was previously married to a Chinese woman and had dated several black women, one for over 16 years, all of whom were unaware of his racist outlook. There was a consistent pattern of, sometimes violent, quarrels with neighbors and he had a track record of obsessive behaviors and grievances. Finally, and most unexpectedly, neighbors regularly witnessed him in his garden either working out or cleaning his car wearing only a thong. The string of Tovey’s behaviors, seemingly inconsequential taken individually, proved an interesting case study for questioning and coding further behavioral material. If nothing else, the project was going to give me some good stories. Tovey’s case was the first we used to train our interns then and remains the first case on which new interns will be trained moving forward. It has that perfect fit of an engaging case study, a rich volume of available variables, and limited noise for coders to sift through.

Ultimately, we formulated over 180 variables in our initial data crawl and pilot. These variables spanned issues concerning socio-demographics, developmental experiences, criminal engagement, violent extremist engagement, pre-attack preparatory behaviors, and attack-related behaviors. While it is true that we could have coded many more variables, we decided to instead prioritize a more limited sampling in the interest of finite timelines and the public availability of certain types of information (e.g. volume and type of extremist content consumed by a lone attacker). Over time, our variable count expanded to over 220, some of which were included to provide greater specificity on features that proved unexpectedly prevalent throughout the cases and required additional expertise to fully understand.

For example, in a large majority of our cases, lone actors were found to have told third parties about their intended attack prior to actually carrying it out, a phenomenon otherwise referred to as leakage in traditional threat assessment. Recognizing this, we decided to gather additional information on the

---

29 For more on Tovey, see: https://www.theguardian.com/uk/2002/oct/04/race.world
30 For in-depth discussion of research processes involving coding media and content consumed by individuals convicted of terrorist crimes, see Donald Holbrook’s RESOLVE Network Chapter: Donald Holbrook, Primary Data and Individual Worldviews: Walking through Research on Terrorist Media Choices (Washington, DC: RESOLVE Network, 2019).
recipient of the leakage, the content of the leakage, how many people the lone attackers told of their violent intent, potential reasons for the leakage, what the recipient did with information from the leakage, and so on.

Similarly, we found a greater than expected presence of mental health problems across lone actor cases. To more fully understand the nature of identified mental health issues, Emily Corner, then a PhD student, joined the team and developed a coding schema for specifying the specific diagnosis, diagnosis date, and associated symptoms. Follow-up projects also necessitated additional information regarding contextual factors, radicalizing settings, and individual propensity, which my colleague and officemate Noémie Bouhanna expertly derived.

Selecting variables of interest and developing a coding schema can be an iterative process, and one that can constantly evolve. Developing and testing a coding schema or codebook on an initial sampling of cases can lead to the inclusion of important variables that may have been overlooked otherwise. In the same vein, consulting experts or literature outside of terrorism studies and/or outside of one’s own initial research team can enrichen and illuminate important new insights from the data itself that are essential to making sense of a phenomenon and the key variables within it.

LESSON 4: CODING DATA & QUALITY CONTROL

Developing a clear and structured coding process is essential to ensuring quality control throughout the data coding process. In order to ensure reliable and consistent coding and minimize missing information, three independent coders coded each observation in our dataset separately. More specifically, for each lone actor, coders filled information into two Word documents. The first was the codebook with the full list of questions (which, upon reflection, really should have been an Excel file to prevent the many headaches the Word document caused later in the project).

The second file contained the original article the coders used to extract responses to the questions in the codebook. If our coders found a relevant piece of information that answered any question on the codebook, they saved the whole original piece, pasted it into the Word file, and highlighted the relevant section in yellow for documentation. Directly after the passage, they indicated which variable(s) the highlighted section pertained to (e.g. Event Variable Number 5). While the process required additional tasks from our coders, in the end, it allowed for much faster and easier re-coding of variables if needed (e.g. mental health, leakage) and built in quality checks (more on this below).

After an observation was coded, the coding results were then reconciled in two subsequent stages. Referring to the three coders involved in the process as coders A, B, and C:

- First, the coding results of coder A were reconciled with the results of coder B.
- Then, the reconciled results from coders A and B were reconciled with the results of coder C.
Cases involving a disagreement between coders were admittedly rare, as coding a “no” response was uncommon. However, in cases where coders could not agree on the correct values for particular variables (e.g. one coder coded “yes” for a variable that another coded “no”), the differences were resolved based on an examination of the comparative reliability and quality of the original sources used by the coders. To assess source reliability and quality, we plotted sources on a continuum of reliability. For example:

**Table 1: Continuum of reliability example**

<table>
<thead>
<tr>
<th>More Reliable</th>
<th>Less Reliable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Court transcripts and associated documents that recorded final judicial decisions were deemed most reliable.</td>
<td>We created a separate continuum for media articles at the less reliable end of the spectrum, considering personal opinion blogs highly subject to personal bias the least reliable and non-tabloid newspapers the most reliable.</td>
</tr>
<tr>
<td>Competency evaluations, sworn affidavits, and indictments, carried out during initial investigations post-arrest and pre-criminal trial were deemed reliable.</td>
<td>Warrants produced prior to an arrest and expert witness reports (in the context of trials) were also considered somewhat reliable. Expert witness reports, in particular, are subject to bias and can be notoriously unreliable.</td>
</tr>
<tr>
<td>Local newspaper reporting of a criminal trial, often much more detail oriented than national outlets, was also deemed reliable.</td>
<td>Verbal or written statements by the terrorist or affiliated group were only deemed somewhat reliable, insomuch as they are particularly susceptible to dishonesty.</td>
</tr>
</tbody>
</table>

Sources at the lower end of the credibility continuum can be highly questionable, subjecting the dataset to potential bias and the possibility of artificially inflating the true prevalence rate (how common the variable was across the entire dataset). Take mental health disorders for example. In many cases, anonymous sources alluded to, but usually did not specify, an individual’s strange and erratic behaviors prior to their radicalization. It is often implied and assumed that this is related to mental health problems, but the reporting is not specific.
For example, on September 23, 2014, Numan Haider stabbed and injured two counterterrorism officers in the parking lot of a police station in Melbourne, Australia. Haider was subsequently shot dead at the scene, but there was evidence that his family previously encouraged him to seek help from a counselor for erratic behavior.\(^{31}\) Their encouragement, however, seemed most likely linked to their worries over his potential radicalization rather than mental health issues.\(^{32}\) Examples like Haider’s fall short of confirming a mental health diagnosis, yet are often regarded as irrefutable evidence of the link between mental health problems and radicalization. To avoid perpetrating potential falsehoods, the bar for confirming evidence should be higher within research and practitioner communities.

Now that empirical evidence has demonstrated lone-actor terrorists have a greater propensity for mental health problems, questions of mental health are increasingly prevalent in media reporting on lone actors. Disturbingly low levels of evidence in a lot of reporting on lone actors depict the actor as having a mental disorder. This may inflate how often mental disorders are reported among lone actors if careful and standardized quality control mechanisms are not in place.

Concerns about over-counting the prevalence rates of certain lone actor characteristics typically stem from the quality of the source, whereas concerns about under-counting prevalence rates stem from issues of critical mass within a coding team and the range of accessible sources available. Not all databases are created equal and those that fail to report basic information about the data within them should be treated with caution. Sometimes studies lack transparency for a reason, sometimes they do so unintentionally. Either way, a lack of transparency can and has resulted in replicability crises elsewhere. Terrorism studies is not immune. Prominent databases similar to our own, even those for which we actually provided our codebook, can be vague on all of these issues and/or have stark data disparities. While differences in definitional choices may impact prevalence rate counts, they are not the only factor at play.

In order to ensure accurate reporting of information and to account for discrepancies in any data presented, increased transparency, attribution, and collaboration in database efforts is needed. This is not only true for lone-actor terrorism database projects, but for all database endeavors. For example, when comparing our data to the Countering Lone Actor Terrorism (CLAT)\(^ {33} \) project—which utilizes a different definition, but similar variables—we identified systematic undercounting of prevalence rates as compared to our own. This includes behaviors like experiencing social isolation (CLAT reported 28% while we reported 52.9%), military training (19% vs. 26%), prior criminal convictions (33% vs. 41.2%), using the internet for tactical research (33% vs. 46.2%), bomb-making manual possession (17.5% vs. 50.4%), and leakage (46% vs. 63.9%). At face value, the reasons behind these disparities are not entirely clear. However, the presence of such disparities is concerning, particularly when it comes to ascertaining the


33 For more, see: https://rusi.org/projects/lone-actor-terrorism
accuracy of any data presented and its subsequent analysis. Greater understanding of the nature and cause of these disparities will ensure greater accuracy in reporting.  

While issues will inevitably arise in any data collection effort, ambitious data collection projects require the requisite amount of foreplanning and institutional capital—be it human, technical, financial—in order to maximize the potential of the volume of qualitative information available.

CONCLUSION

No matter what the topic, research has the potential to impact policy and practice. In terrorism studies and related fields, this should not be taken lightly. I have seen our lone-actor work cited to justify portions of thinking behind particular interventions, including the establishment of a joint agency response to the threat of lone-actor grievance-fueled violence in Australia,  


While issues will inevitably arise in any data collection effort, ambitious data collection projects require the requisite amount of foreplanning and institutional capital—be it human, technical, financial—in order to maximize the potential of the volume of qualitative information available.


41 For more on TRAP-18, see: https://www.gifrinc.com/trap-18-manual/. For more on VERA2R, see: https://www.vera-2r.nl/.

the research is responsibly conducted, accurate, and trustworthy, but also that it contains the necessary
validity and caveats necessary to be used responsibly in policy and practice over time.43

I have seen our lone-actor research miscited to justify policy positions I find questionable. Whilst my
usual reaction might be a Twitter rant, such instances always makes me question whether I could have communicated the research more effectively and clearly. Should I have taken more time with the abstract? Should I have put the caveats up front rather than burying them in the conclusion? Should I publish non-technical blogs to coincide with the journal article? The answer to effective translation of science into policy remains elusive. Despite this, what remains relevant, and what this chapter intends to communicate, is the importance of understanding and taking into consideration how data constraints, research processes, and the choices of a research team throughout the research process impact the strength of research, necessitate transparency around all aspects of the research process, and carry significant implications for the use of research in policy and practice.

**SOURCES**


The RESOLVE Network is a global consortium of researchers and research organizations committed to delivering fresh insight into violent extremism around the world. The Network provides access to open-source data, tools, and curated research to ensure policy responses to violent extremism are evidence based. Members of the Network work in parts of Africa, Asia, Europe, and the Middle East to promote empirically driven, locally defined responses to conflict and to support grassroots research leadership on violent extremism.

Our partners operate in more than 25 countries where challenges with conflict are an everyday reality. We are passionate about amplifying credible local voices in the fight to mitigate the destabilizing risks of social polarization and political violence. The RESOLVE Network Secretariat is housed at the U.S. Institute of Peace, building upon the Institute’s decades-long legacy of deep engagement in conflict-affected communities.

To learn more about the RESOLVE Network, our partners and how to get involved visit our website, www.resolvenet.org, and follow us on Twitter: @resolvenet.