

Exploring the Landscape of Sentencing for Terrorist Offenders:

A Scoping Review

The field of counter-terrorism studies has witnessed significant growth over the past two decades. Despite this growth, research on the sentencing of terrorist offenders remains relatively limited. This scoping review aims to comprehensively map this research area in terms of volume, nature, and characteristics, adopting a rigorous and transparent approach, resulting in the inclusion of 59 studies. After mapping these studies according to their general characteristics (specifically the historical and geographical distribution, and the adopted research methods), this review delves deeper into the specific characteristics of quantitative studies (including the timeframe of each dataset, the sample sizes, the coded variables and the statistical strategies used). A thematic analysis of the studies reveals key findings concerning legislation and its impact on terrorism trials, sentencing principles and guidelines, as well as specific factors influencing sentencing outcomes. Additionally, the review features recommendations for future research, as outlined in existing studies. The discussion section outlines the implications for policy and practice, presents recommendations for scholars, and reflects on the limitations of this review. In particular, a plea is made for more mixed methods-driven research within European countries, by establishing open-source databases based on official documentation.

Keywords: terrorism; terrorist offender; criminal justice system; sentencing; sentence outcome; judiciary; scoping review

Introduction

To date, we know relatively little on the interaction between the criminal justice system, terrorist offenders, and the influences of changes in legal and/or political context (Bradley-Engen et al., 2009). An initial assessment of the existing body of literature shows that the research area focusing on sentencing terrorist offenders is underexposed in the vast research domain of (counter)terrorism. To illustrate the scope of the research domain, a Web of Science search on the topic of “terrorism” results into over 26 000 hits,

with over 90% published after 9/11 (despite the first hit dating back to 1959). In contrast, the same search tactic combined with the keyword “sentencing” produces merely 158 results over a timespan of 33 years (1989-2022), with almost all publications (94%) appearing after 2001. Existing studies on the topic corroborate that criminal justice responses to terrorism, and sentencing in particular, remain a largely underexposed issue in terrorism studies (Amirault & Bouchard, 2017; Amirault et al., 2016; Bradley-Engen et al., 2009; Diab, 2013b; Johnson, 2012; Said, 2014; see also Yon & Milton, 2021).

Yet, in evaluating the criminal law as a response to terrorism, it is key to study the actual sentencing decisions of judges (Roach, 2014). Presenting research findings on sentencing terrorist offenders, with a focus on the trial phase and sentencing outcomes, holds significant relevance. This approach addresses a substantial knowledge gap, contributes to a more holistic understanding of counter-terrorism efforts, informs policy and practice, and influences sentencing practices. Especially given the unique dimensions of terrorism-related crimes and offenders (Amirault & Bouchard, 2017; Pyne, 2011), understanding how terrorist offenders are sentenced and which factors influence these sentencing outcomes is pressing. Due to the political nature of these offences, the judicial decision-making processes, sentencing principles and key predictors differ from those in the context of traditional crimes (Istiqomah, 2020). Gaining insight into these processes, principles and predictors is therefore essential for both policymakers and practitioners, as well as scholars. Hence, the goal of this scoping review is twofold: (1) to provide guidance for policymakers (in creating an apt legal framework) and practitioners (especially prosecutors, in seeking legitimate convictions), and (2) to offer insights and recommendations to scholars in the field of counter-terrorism, thereby contributing to the development of future research that can, in turn, provide valuable input for policymakers and practitioners. Furthermore, the evolving nature of terrorism threats has presented

significant challenges for all stakeholders involved, leading to an unprecedented expansion of the preventive paradigm and the use of reactive and repressive criminal law measures for pre-emptive purposes (see De Coensel, 2020; Van der Woude, 2010).

It is, therefore, highly needed to examine how this field of interest has been approached in scientific literature (in terms of volume, nature and characteristics of the studies). This article aims to provide a comprehensive overview of this research area, clarify key concepts, examine similarities and differences across studies, and identify research gaps. While the primary focus is aimed at judicial decisions in terrorism proceedings, encompassing aspects of guilt, interpretation of constitutive elements, sentence type and sentence length (in other words, sentencing outcomes), the thematic analysis reveals a wide and diverse range of research findings within the included studies on sentencing terrorist offenders.

The article will firstly explain the methodology of this scoping review and report on its different stages. Thereafter, the results section will present the data in terms of volume, nature and characteristics and summarize the evidence available. Finally, the discussion and conclusion will link the main results to the research objective of this article, address the limitations of this scoping review process and formulate future research recommendations.

Methods: A Scoping Review

Today, literature reviews come in all shapes and sizes (see Arksey & O'Malley, 2005; Grant & Booth, 2009). A scoping review is a relatively new type of review (Munn et al., 2018), which can be defined as “a form of knowledge synthesis that addresses an exploratory research question aimed at mapping key concepts, types of evidence, and gaps in research related to a defined area or field by systematically searching, selecting, and synthesizing existing knowledge” (Colquhoun et al., 2014). This ‘evidence synthesis

approach' has different objectives compared to a traditional systematic review (Munn et al., 2018; Pham et al., 2014), by focusing on broader topics and different study designs, while the quality of included studies is not necessarily assessed (Arksey & O'Malley, 2005; see also Munn et al., 2018; Peters et al., 2020). This method is therefore the preferred tool "to determine the scope or coverage of a body of literature on a given topic and give clear indication of the volume of literature and studies available as well as an overview (broad or detailed) of its focus" (Arksey & O'Malley, 2005; Munn et al., 2018; Peters et al., 2020) – which this article aims to conduct in the field of sentencing terrorist offenders.

This scoping review is based on the influential framework proposed by Arksey and O'Malley (2005), while incorporating the enhancements recommended by Levac et al. (2010) and Peters et al. (2020). Additionally, it follows the preferred reporting items of the PRISMA-ScR checklist¹. Any deviations from this framework are transparently reported. The following subsections align with the five² framework stages of Arksey & O'Malley (2005): (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarising and reporting the results.

Framework Stage 1: Identifying The Research Question

In this article, the aim is to explore the research landscape concerning the sentencing

¹ PRIMA, short for Preferred Reporting Items for Systematic Reviews and Meta-analyses, is a widely recognized evidence-based minimum set of items to help authors improve the reporting of systematic reviews and meta-analyses. Since 2018, the PRISMA Statement is extended to scoping reviews with the PRISMA-ScR checklist. See <http://www.prisma-statement.org>.

² The sixth and optional stage of consultation was not conducted in this scoping review.

outcomes of terrorist offenders by addressing the following central research question: “What are the key research trends and methodologies employed in the scientific literature regarding the sentencing of terrorist offenders, and how have these evolved over time in terms of volume, nature, and characteristics?”. Each of these parameters must be defined at the outset of the scoping study:

- Sentencing: This research focuses on the trial phase of the criminal justice processing of terrorist offenders and aims to shed light on judicial decisions, including aspects of guilt, interpretation of constitutive elements, sentence type and sentence length (together summarized as ‘sentencing outcomes’). Instead of restricting this review to the imposition of the penalty (post-conviction phase), a broader approach is adopted.
- Terrorist offenders: Terrorist offenders are defined broadly in this review. This definition includes not only those who committed core terrorist offences but also those involved in offences related to a terrorist group or to terrorist activities (including incitement, travelling for terrorist purposes, receiving or providing training, recruitment, material support, or similar common charges).
- Volume: The volume of scientific literature relates to the quantity and distribution in time.
- Nature: The nature of scientific literature refers to the type of studies, in terms of scientific field and methodologies used.
- Characteristics: The characteristics of scientific literature concern the territorial scope of the study, timeframe, sample set, parameters or themes, and key findings.

Framework Stage 2: Identifying Relevant Studies

The identification of research studies must be as comprehensive as possible (Arksey &

O'Malley 2005). Due to practical time and budget constraints, two key decisions have been made on the coverage of the review in terms of language and format. Selected research studies had to be written in English, but publication status was less restrictive. This allowed us to include grey literature (such as reports, working papers, and PhD dissertations) to maximize the comprehensiveness of the review. Bachelor theses, Master theses, speeches and newspaper articles, however, were excluded from the analysis. No delineation was made based on the year of publication, since the preliminary view on the existing body of literature had shown that this was unnecessary to manage the sample.

The records were identified through a search of the electronic databases and the hand-searching of reference lists. In total, six databases were searched: BASE (Bielefeld Academic Search Engine), Proquest, Science Direct, Scopus, Web of Science and Wiley Online – with a supplemental check of Google Scholar. These databases were carefully selected based on the work of Gusenbauer & Haddaway (2020). In their article on the evaluation of 28 academic search systems for systematic reviews or meta-analyses, search systems were either rated as a principal or supplementary resource as a result of 27 test procedures and performance requirements. For the purposes of this review, only *multidisciplinary* databases rated as *principal* were selected. Google Scholar was added as a supplementary resource, especially in the search for grey literature. However, only the first 100 hits (of the over 190 000 hits) were inspected, since “[Google Scholar] is more concerned with ‘tuning’ its first results page than with overall precision [making it] highly precise for exploratory searches conducted by a user interested in only a few relevant results on the first search engine results page” (Gusenbauer & Haddaway, 2020).

The search strategy was multiple times adapted throughout the process. A balance had to be sought between efficiency, effectivity, and the level of precision. After piloting the search strategy, the final research strategy included following search terms for the

population, the concept, and the context³: (terrorism OR terrorist*) AND (sentencing) AND (charge OR convict* OR trial OR "criminal justice" OR judicial). This search was updated until December 1, 2022. Endnote was used as a valuable data management tool to keep track of the identified reports that were included based on title or abstract. As a result of the fact that the first-stage screening took place in the online interface, the exact number of duplicates was not registered. Of the included reports based on title and abstract, 17 reports showed multiple times within and throughout the different databases. The 370 records were spread across the different databases as follows: BASE n=119; ProQuest n=93; Science Direct n=8; Scopus n=68; Web of Science n=69; and Wiley Online n=13. 100 additional hits were assessed on Google Scholar.

Framework Stage 3: Study Selection

The cumulative inclusion criteria used in this scoping review relate to (a) the focus on the offending population of terrorist offenders, (b) the focus on sentencing outcomes, and (c) the delineations made based on language and format. Although the search terms clearly focused on the offending population of terrorist offenders and sentencing outcomes, some records either focused on broader topics related to terrorism (for example, incarcerating terrorist offenders, prison strategies, deradicalization, post-release supervision, recidivism, risk assessments, public reactions, the legislative framework, or profiling), or on sentencing studies in general (without a focus on the context of terrorism-related offences). Even though some of these studies might include interesting data (for example in the context of recidivism studies or studies on foreign fighters), solely records that explicitly focus (at least in part) on the sentencing stage of terrorist offenders are included.

³ Peters et al. (2020) have introduced these threefold 'Population', 'Concept' and 'Context' elements to structure the identification of relevant studies.

To ensure proper delineation of the review, at least one research question or section had to relate to the sentencing of terrorist offenders (see Berkell, 2017; Harms, 2017; Heinrich et al., 2012; Horgan et al., 2018; Smith & Damphousse, 2002)). When the database search resulted in an entire edited book, this single report was broken down into multiple, relevant book chapters. Likewise, when a relevant book chapter was identified, other chapters within the edited book were searched for interesting reports. The edited book itself is not included as a separate study.⁴

First, all 470 records were screened on the basis of title, abstract and/or table of content (first-stage screening). Afterwards, the remaining reports (n=69) were assessed on their eligibility by reading the full text (second-stage screening). A list of reports that were excluded in this second-stage screening is provided in an annex. By hand-searching the reference lists of the included studies (n=51) and using previously gathered sources, 8 additional relevant studies could be included. In total, 59 studies are included in the review. The entire process of identification, screening, eligibility and inclusion is reported through a flow diagram (Figure 1) (Moher et al., 2009). The included studies were not assessed on their quality of evidence (no methodological appraisal).

(INSERT FIGURE 1)

Framework Stage 4: Charting the Data

In a next stage, the data was synthesized and interpreted “by sifting, charting and sorting material according to key issues and themes” (Arksey & O’Malley 2005). The data were

⁴ As was the case for De Graaf, B., & Schmid, A. P. (2016). *Terrorists on Trial: A Performative Perspective*. Leiden University Press; Loadenthal, M. (2021). *Prosecuting Political Violence: Collaborative Research and Method*. Routledge; Nesbitt, M., Roach, K., Hofmann, D., & Lee, K. (2021). Canadian Terror: Multi-Disciplinary Perspectives on the Toronto 18 Terrorism Trials. *Manitoba Law Journal*, 44(1).

charted in Excel by recording general and specific information about the study (including author, year of publication, title, key words, aims of the study, study design/methodology, geographic scope, timeframe, sample size, coded variables, main research findings and recommendations future research). This ‘descriptive-analytical’ method (Arksey and O’Malley 2005) results in a common framework to collect and report all standard information on each study.

Framework Stage 5: Collating, Summarising and Reporting the Results

The main working document consisted of an Excel file, in which the charted data was further coded and analysed. By applying a consistent approach in both the charting, coding, and reporting phase, the author remains as neutral as possible towards the included studies and presents a comprehensive and thorough review of the available literature.

On the one hand, a numerical analysis was conducted to map and code the studies, utilising Excel’s filtering and pivot-table functions. Tables and charts support this mapping exercise to illustrate the historical and geographical distribution of the studies, and the characteristics of quantitative studies in term of timeframe, sample size, data sources and coded variables. PowerBI was a helpful tool in data visualisation. On the other hand, all included studies were organized thematically to highlight the different elements and focal points in the literature. The thematic categorisation is the result of a dynamic and inductive coding process, without a pre-arranged codebook. This thematic analysis will not only focus on the main research findings, but also on the recommendations for future research as formulated by the authors.

Results

General Characteristics

Evolutions over Time: Authors and Geographic Scope. The search strategy of the scoping review did not make any delineations regarding to the year of publication. However, the first relevant study aligning with the inclusion criteria was published in 1996 (Smith & Damphousse, 1996). Figure 2 shows the dispersion over time (with 1996 as a starting point) and highlights that approximately over 80% of all included studies were published since 2010.

INSERT FIGURE 2

More than half of the included studies exclusively examine data from the United States. Research outside of the United States began to gain traction since 2010. Leading countries in non-US studies are Canada (15.25%), the United Kingdom (8.47%), Australia (5.08%), the Netherlands (3.39%), and studies that involve a combination of these countries in comparative analyses (8.47%) (see Figure 3). Single studies on Germany, Norway and Indonesia are also identified.

INSERT FIGURE 3

Evolutions over time: methods. When we look at the methods used over time, we see that a qualitative approach gained significant weight since 2010 (55.93%). These qualitative studies are diverse in nature and contain literature studies, legal analyses, or anecdotal evidence. Some qualitative studies focus on one case in particular (7 out of 33 qualitative studies: Brown (2012); Diab (2013a); Hemmingby and Bjørge (2016); Hoffmann (2019); McCarthy (2004); Nesbitt (2021a); Pyne (2011)), whilst other studies aim to map the jurisprudence on terrorist offenders in a more general manner. Often, the methodologies used in these qualitative studies are not transparently reported on. Quantitative methods

(38.98%) are quite evenly distributed over time. A distinction is drawn between descriptive and inferential methods. This distinction categorises studies that solely summarise dataset characteristics as descriptive, while differentiating them from studies that enable hypothesis testing or the assessment of data generalisability to a broader population, which are categorised as inferential. Coding this traditional distinction proved to be less straightforward in practice. The author opted to code every study that makes use of a significance test as ‘inferential’ (for example crosstabs with Pearson chi-square test) and not only those studies that perform more complex statistical strategies (such as regression or ANOVA). It is worth noting that one study did not rely on real-life sentencing data, but has created a simulation through an online survey (Frings et al., 2018). Mixed methods were relatively rare and recent, with only three studies out of 59 falling into this category, all published in 2019 and 2020.

Authors & disciplines. The entire dataset includes a total of 98 authors, with 69 of them being unique contributors. Notably, during the initial years of research in this field, three authors, namely Kelly R. Damphousse, Brent L. Smith, and Chris Shields, published extensively on the topic. To gain insights into the disciplinary backgrounds of all authors involved, their educational and professional backgrounds was considered, with a specific focus on their field of doctoral studies. As a result, a majority of the included studies can be categorised within the legal domain, accounting for 54.24% of the total. Other prominent disciplines represented include sociology or social sciences with an emphasis on crime and justice (20.34%), followed by criminology (8.47%). Additionally, less prevalent disciplines were identified such as historical sciences, psychology, political sciences, law and economics, as well as engineering sciences.

Theoretical framework. In the literature, relatively little attention is paid to theory testing or theory development. Notable exceptions are authors Kelly R. Damphousse, Brent L. Smith and Chris Shields, who have consistently used (a combination of) the following theories within their work: consensus theory, conflict theory and structural-contextual theory (coupled with the liberation hypothesis and the trickle-up perspective) (Smith & Damphousse, 1996; Smith & Damphousse, 1998; Shields, Damphousse & Smith, 2006; Damphousse & Shields, 2007; Bradley-Engen, Damphousse & Smith, 2009). These theories relate to the relationship between legal versus extralegal variables and their impact on sentencing outcomes. Consensus theory posits that sentencing is based on a societal consensus, implying that legal variables are the primary predictors of sentence severity. In contrast, conflict theory suggests that sentencing decisions may be influenced by power dynamics, social inequalities, and conflicts of interest, leading to a more significant role for extra-legal variables. Structural-contextual theory proposes that the predictability of sentence outcomes varies for different types of crimes. The conflict perspective is the only theory that was also more recently tested by Amirault & Bouchard (2017). Other theories or broader relevant concepts that came across in the literature were the Duffian theory of punishment (Diab, 2013); secondary risk management, anticipatory prosecution and actuarial justice (De Graaf, 2016; De Graaf, 2019); and cognitive behavioural theory, framing effects and cognitive biases (Nesbitt, Oxoby & Potier, 2019). In contrast to these deductive research studies, three studies provide an application of grounded theory (Burtis & Butler, 2021 on developing sentencing categories from a governmental perspective; Bielamowicz, 2021 on identifying the consequence of perceived foreignness on the prosecution of political violence cases in the United States; Weaver & Doty, 2021 on gender interaction effects).

Quantitative Studies: A Closer Look

In the previous subsection on the general characteristics of the included studies, it became clear that almost 39% of the studies are of a quantitative nature, and only 5% make use of mixed methods. To grasp the characteristics of these studies more in-depth, this section will map the timeframe of each dataset, the sample sizes, the coded variables and the statistical strategies used (see Table 1 for a summary on each study).

INSERT TABLE 1

Timeframe. Two types of codes were created to gain insight into the timeframe of the datasets of the included studies: whether the comprised data concerned the pre- or post-9/11 era on the one hand, and the duration of the timeframe on the other hand. The rationale for distinguishing between pre- and post-9/11 data is prompted by the general acknowledgment that 9/11 has triggered new anti-terrorism laws which “changed the sentencing paradigm from post-crime criminal justice to pre-crime national security” (Istiqomah, 2020), in addition to the fact that “prior to 2001, there were relatively few data-driven studies” in this regard (Gruenewald et al., 2022). The findings indicate that 38.46% of the included studies exclusively use data from the post-9/11 era. Over 30% of the studies used data from both the pre- and post-9/11 era, which was also the case for studies that focused solely on pre-9/11 data.⁵

Almost 27% of quantitative studies entailed a very broad timeframe of over 20 years. The same goes for the category of 16 to 20 years of data. Over 34% analysed

⁵ When there was only a one-year difference with 9/11 as the baseline, this was not taken into account. Studies with a timeframe of 1987-2002 for example, were coded as pre-9/11, whilst a timeframe of 2000-2007 was coded as post-9/11.

between 6 and 10 years of data. The remaining forks of 11-15 years and 1-5 years of data came across only once.

Sample size. In summarising the sample size of the datasets, only the number of terrorist offenders was considered. If datasets included non-terrorist offenders for the purpose of comparison, this information was deemed irrelevant for the current analysis. Over 38% of the quantitative studies used a dataset comprising 100 to 200 terrorist suspects, which is the same percentage as studies that had larger datasets of more than 200 terrorist offenders. Over 23% of the quantitative studies employed smaller datasets of under 100 terrorist defendants.

Data sources. In the early studies, a major drawback of research was highlighted, namely the lack of high-quality, empirical, national-level data available for analysis (Smith & Damphousse, 2002). The Terrorism Research Center at the University of Arkansas has, therefore, created an empirical database (namely the American Terrorism Study; ATS), based on the US Federal Bureau of Investigation's definition of terrorism, "from which criminological theories and governmental policies could be effectively evaluated" (Smith & Damphousse, 2002). Today, the dataset is labelled as the "longest running project on domestic terrorism in the United States" (Gruenewald et al., 2022). The dataset consists of data on federal terrorism-related court cases, persons indicted in these court cases, and related officially designated terrorism incidents, and is divided into five sub datasets on (1) 'counts data', (2) 'indictees data', (3) 'persons data', (4) 'cases data', and (5) 'group data'. All cases are collected using the Public Access to Court Electronic Records (PACER) system. The founders believe that the creation of an empirical database overcomes deficiencies related to the collection of empirical evidence and the use of varying definitions of terrorism, since the data comprised in the database are acquired by

releases by the FBI of lists of persons indicted as a result of investigation under the FBI's Counterterrorism Program. After receiving such a list, the cases were reviewed at the federal district court or at the federal regional records center. Many included studies rely on the American Terrorism Study (10 out of 26 quantitative or mixed methods studies).

Today, the ATS is integrated in the 'Terrorism and Extremist Violence in the United States Database' (TEVUS), together with three other open-source datasets, namely the 'Global Terrorism Database' (GTD), the 'US Extremist Crime Database' (ECDB), and the 'Profiles of Perpetrators of Terrorism in the United States Database' (PPT-US). In the included studies, there is – in addition to the ATS – only a reference to the ECDB (Gruenewald et al., 2022).

Other datasets that were referred to in the included studies are the Prosecution Project Dataset (based on Open-Source Intelligence) (Burtis & Butler, 2021; Bielałowicz, 2021; Weaver & Doty, 2021), the Profiles of Individual Radicalization in the United States (PIRUS) dataset (Yon & Milton, 2021) and the Officially Adjudicated Terrorists in Canada (OATC) dataset (Amirault and colleagues 2015, 2016, 2017). Open-Source Intelligence refers to the process of finding cases through a daily review of Twitter feeds and automated news alerts and the manual scraping of "extensive reports, tables, indexes, datasets, and other large collections" (Loadenthal, 2021). When no reference to a dataset was made, the studies were based directly on official data (including publicly available sentencing decisions or legal documents provided by the police, prosecution services or courts) or open source data (more in particular, news reports).

Coded variables. A scoping review of literature on sentencing terrorist offenders should not only shed light on the general characteristics, research methods, and highlighted themes but should also delve into the coded variables in quantitative studies. These variables not only reveal the focus of existing research but also highlight areas that have

received less attention. This list of coded variables could serve as a valuable reference for creating a codebook in future research. Many US studies are based on data of the American Terrorism Study, which includes approximately 80 variables that measure demographic information on the defendant, general case information and terrorism-specific information. While most studies use the individual as the unit of analysis, data is also collected on count level (see Harms, 2017). The overwhelming majority of quantitative studies code *sentence length* as the dependent variable (measured in months). Yon & Milton (2021) have pointed out that while sentence length is “an important measure of the outcome of an investigation, [it] is only one of many measures where differences may manifest”. These authors have therefore incorporated *alternative outcome measures* into their analysis to examine the concept of severity (such as the decision to pursue investigations and the type of charges made). In exceptional cases, it is rather the *sentence type* instead of the sentence length that is of importance (one study even specifically measured the likelihood of receiving incarceration in addition to the length of incarceration, see Bradley-Engen et al. (2009)). More common, is the case outcome in terms of *disposition mode* (guilty plea versus trial conviction).

Frequently occurring independent variables relate to the *type of offense/charges* (sometimes making a distinction between terrorism-specific provisions, common criminal law provisions or a combination of both; at other times making more thematic distinctions such as weapon charges versus racketeering charges versus violence charges, etc.), to the *offence/injury severity* (measured by the level of threat to human injury or by the degree of punishment), to the *number of counts/charges* (which is often seen as an indication of the complicated nature of the case), to the *type of terrorism* (making a distinction between jihadist, right-wing, left-wing, ethno-nationalist and separatist, and single-issue terrorism, or other distinctions such as domestic versus international

terrorism), to the *role of the offender* (lone wolf versus leader of a terrorist group versus subordinate) and to the *number of co-defendants*.

In most quantitative studies, sociodemographic variables were used as control variables. These control variables most often referred to the *age*, *gender*, *ethnicity*, and *criminal history* of the individual, followed by *education*. Variables such as *citizenship*, *military background*, *religiosity* and ‘*otheredness*’ were considerably less frequently observed. Yon & Milton have created ‘*biographic availability*’ as a summarising dummy variable, accounting for indicators “that might signal non-existent or reduced connections in society that might encourage someone to engage in higher risk behavior, or that they might have few connections to help with rehabilitation and disengagement”. These indicators relate to relationship status, parenthood, education, and employment.

Variables that returned in less than six studies relate to:

- (1) Other procedural elements: including *prosecutorial* and *defense strategies*, the *type of attorney*, variables on *bail*, and *court characteristics* – in the sense of location of the court, terrorism case load, and terrorism trial rate;
- (2) Other time-related elements: including the *date of indictment*, the *timing of the offense* or the *timing of arrest* in order to study the effect of incident-based contextual factors on sentencing outcomes; pinpointing the timing regarding the *pre- or post-guideline era*; the *months to conviction* and the *time spent in pre-trial detention* since the duration of the case processing;
- (3) Other offence-specific elements: including *political motivation* (and thus making the distinction between terrorist and non-terrorist offenders), *type of weapon* and *use of firearms*, presence of *aggravated circumstances*, *completion* of the offence, and the *location* of the preparation and/or attack;

- (4) Other group-related elements: including the precise *terrorist organization*, the intended and actual *targets* of the group, and the *identification number* of a certain group or event to represent associations among the accused;
- (5) Other evidence-based elements, in terms of the *type of evidence* and the *number of items of evidence*; and
- (6) Other behaviour-specific elements, including a wide variety of variables that map the *behavioural functions* of defendants, for example in terms of material support, online activity or the attendance of a training camp.

Statistical strategies. The statistical strategies used to study the sentencing of terrorist offenders range from descriptive statistics to bivariate analyses and multivariate modelling methods.

While some studies provide a descriptive summary of the features of the collected data, others add statistical tests to their analysis for the purposes of testing whether there is a statistical relationship between two variables (correlation coefficient or chi square), or whether two groups or variables are statistically different from each other (t-test or chi square, depending on the measurement level of the dependent variable).

When it comes to multivariate modelling techniques, most studies use multiple linear regression models (more in particular, Ordinary Least Squares (OLS), see Amirault and Bouchard (2015); Amirault et al. (2016); Bradley-Engen et al. (2009); Bradley-Engen et al. (2012); Heinrich et al. (2012); Istiqomah (2020); Smith and Damphousse (1998); Yon & Milton (2021)) and logistic regression models when the dependent variable is nominal or dichotomous in nature (see Frings et al. (2018); Harms (2017); Heinrich et al. (2012); Shields et al. (2006)) in order to observe the impact of various variables on the odds of conviction and sentence length. More complex strategies used are:

- (1) Structural Equation Modeling (SEM), which is similar to regression analysis, but more powerful since SEM allows to develop complex path models with direct and indirect effects (see Smith & Damphousse, 1996; Smith & Damphousse, 1998);
- (2) Multilevel Modeling strategies (MLM), which account for interdependencies in the data. Johnson (2012) is a strong opponent of this useful, but underutilised, approach in the context of terrorism studies since terrorist offenders are nested within multiple macro-social contexts (such as courts and terrorist networks). This technique has been applied in the studies of Johnson (2012) and Amirault and Bouchard (2017), whilst in another study the need for a hierarchical linear model was at least tested (but was deemed not necessary, since the intraclass correlation was not significant: see Amirault et al. (2016));
- (3) Multidimensional Scaling methods (MDS) to calculate the likelihood that two variables (or a series of variables) will co-occur in the same case or person (see Horgan et al., 2018);
- (4) Timeseries plots (Amirault et al., 2016), although this method was not explained in further detail;
- (5) Corpus linguistic analysis in R to examine word choices or lexical patterns, with a view on determining the rhetoric of sentencing hearings, sentencing memorandums, judgements and transcripts of court proceedings (Weaver & Doty, 2021).

Thematic Analysis

The foregoing sections have shed light on the ‘measurable’ characteristics of the included studies. This section, however, will delve into the thematic foci within studies on sentencing terrorist offenders. After a thorough reading of the included studies, a twofold classification is made, namely themes relating to (1) legislation, trials, sentencing

principles and guidelines, and (2) factors influencing sentencing outcomes. As already mentioned, the thematic categorisation is the result of a dynamic and inductive coding process and demonstrates the rich array of diverse research findings. At the end of this section, calls for future research are summarised.

Legislation, Trials, Sentencing Principles and Guidelines

Given that this article focuses on the trial phase of the criminal justice processing of terrorist offenders, this review goes beyond the mere imposition of the penalty. Therefore, before turning to the specific factors that influence the sentence outcome, an analysis is made of elements regarding legislation and the function of terrorism trials. The main emphasis will be placed upon the ratio behind sentencing and sentencing principles.

Legislation. In the post-9/11 era, counter-terrorism measures with a preventative and proactive outlook proliferated. As a result, new types of terrorism cases are prosecuted in which risk management plays a central role, also referred to as ‘anticipatory prosecution’ (De Goede & De Graaf, 2013; De Graaf, 2016; see also Istiqomah, 2020). Lee and Walker (2022) focus particularly on the sentencing of so-called precursor crimes, arguing that these crimes do not only “represent manifold challenges to the legitimacy of the criminal law and criminal process”, but that these challenges also persist in the phase of sentencing. As Diab (2013a) noted: “jurisdictions had developed a jurisprudence on terrorism sentencing in cases ranging from principal figures in very serious plots to peripheral figures with minor or tangential roles”.

While one of the objectives of terrorism-specific legislation is to increase sentences for these offenders, research has established that offenders sanctioned under general legislation, or a combination of general and terrorism-specific offences, are punished more severely (Amirault & Bouchard, 2017; Amirault et al., 2016). Research

has found that, following the implementation of terrorism-specific legislation in Canada, the average sentence length of terrorism-related offenders decreased and, therefore, suggests that this decrease “is perhaps better characterized by the relative success of law enforcement in disrupting terrorist plots rather than by a failure in the new legislation to achieve harsher punishment” (Amirault et al., 2016). As such, the rate of offenders who are able to successfully complete an offence decreased (Amirault et al., 2016). In Canada, non-terrorism legislation is still most consistently utilized in terrorism cases (Amirault & Bouchard, 2017). Nesbitt and Hagg (2020) have judicially interpreted the Canadian provisions through an empirical analysis of court decisions. This contribution is one of the few studies within this scoping review which focuses upon the judicial interpretation of the constitutive elements of terrorist crimes. The authors criticize the overlap between the offences, the overbroad drafting and the response of the judiciary, who has chosen to read down the provisions rather than overturning them as unconstitutional (Nesbitt & Hagg, 2020).

Trials. De Graaf (2016) claims that “terrorism trials are key opportunities to legitimise the scope and substance of post-9/11 terrorism legislation, which is simultaneously implemented, contested and performed.” De Graaf (2016, 2019) witnesses the transformation of terrorism trials from a precautionary or actuarial justice perspective and considers these trials as a performative space. She points out that the virtual or premeditative turn in the criminal justice system, through the incorporation of precautionary measures into criminal legislation (as discussed above), “is not just a matter of juridical change, but plays out in the performative dynamic between the police, the public prosecutor, defence counsel and the judges” (De Graaf, 2016). The courtroom has become the place where trajectories of radicalization are invoked in order to prevent acquittals (De Graaf, 2019). By sentencing potential violence, terrorism trials can be

understood as tools of secondary risk management. Whereas De Graaf (2016, 2019) focuses on terrorism trials as performative spaces, especially from a prosecutorial and judicial perspective, Hemmingby and Bjørgo (2016) witness the same trend from the defendant's perspective. Their research focuses on the Breivik case, in which they conclude that the defendant played "the court as a stage" and had a "key role with an opportunity to present his ideology and views to a wider audience". These contributions are prime examples of literature on terrorism trials that focus on one high-profile case, such as the Hofstad Group in the Netherlands (De Graaf, 2016), the Breivik case in Norway (Hemmingby & Bjørgo 2016) and the Toronto 18 in Canada (Nesbitt, 2021a).

McCarthy (2004) represents a notable voice in the debate, arguing that when examining the effectiveness of the criminal justice system, particularly in the context of terrorist offenders, a distinction must be made between a due process versus a national security / public safety perspective. In his article, McCarthy (2004) discusses Al-Qaeda trials and, from a due process perspective, asserts that the rule of law is fully adhered to. However, he argues that the principal concern should lie with public safety considerations. In that regard, he concludes that criminal trials are not suitable for addressing terrorism, advocating for a comprehensive counterterrorism strategy that should include "all of the tools of government: military, diplomatic, intelligence, financial, and to a far more limited extent than throughout the 1990's, law enforcement".

Judicial reasoning & sentencing principles. The majority of qualitative studies focuses on the process of judicial reasoning and sentencing principles (deterrence, rehabilitation, denunciation) in terrorism-related cases (Berkell, 2017; Crowley, 2010; De Graaf, 2019; Diab, 2011, 2013a; Lee & Walker, 2022; McGarrity, 2013; Nesbitt, 2021a; Nesbitt et al., 2019; Pyne, 2011; Roach, 2014; Said, 2014; Sameer, 2016; Scanlon, 2014; Zaia, 2017). In general, these contributions almost all denounce the emphasis on deterrence in

sentencing terrorist offenders, while considerations of rehabilitation are scarcely mentioned. As Diab (2013a) witnessed in his case law analysis: “terrorism offences were a special category of crime and called for a primary, if not exclusive, emphasis on denunciation, deterrence and incapacitation”. Crowley (2010) criticizes this “being tough on terrorism” policy, leading to the erosion of fundamental rights, a utilitarian system of general deterrence and the empty promises of rehabilitation. De Graaf (2019) refers to the shift “towards control and deterrence rather than focusing on re-integration, rehabilitation and transformation of the convicted offender”. Focus on rehabilitation also presents a central theme within the contribution of Sameer (2016), in which he draws lessons from the War on Drugs. Whereas the War on Drugs and the War on Terror have both resulted in lengthy punitive incarceration (often of young Muslims and African Americans), policymakers have instituted reforms to reduce the length of drug related sentences and to focus on alternative measures and rehabilitation. As Sameer (2016) argues, these insights have not yet been translated to the War on Terror. Pyne (2011) remarks that the focus has changed from the harm caused to the challenges presented by terrorist ideology. McGarrity (2013) takes a more moderate stance, by stating that “the overwhelming determinant of the sentence handed down in a terrorism case is the objective seriousness of the offence” – even though all terrorism offences are considered serious. However, according to the analysis of McGarrity (2013), “courts take into account the target of the planned terrorist act, the damage that the terrorists intend to cause and the proximity between the preparatory acts and the commission of a terrorist act”. Scanlon (2014), on the other hand, stresses that “greater consideration of proximity in sentencing processes may lead to sentences that are more proportionate to the gravity of the offender’s conduct and intention”. Roach (2014) pleads for a case-by-case approach, in which offender characteristics (such as youth, sophistication, future danger and amenability to

rehabilitation) should be considered, instead of focusing merely on the terrorist nature of the offence. Nesbitt (2021) comes to the same conclusion as he argues that aspects of individuality are downplayed “in favour of a generalized assessment of the seriousness of terrorism in general”. As a result, the normal balance between individual moral culpability and the seriousness of the offence is skewed (Nesbitt et al., 2019; Nesbitt, 2021). Diab (2011) expresses a more exceptional stand of view, given that he advocates for greater discretion to judges in the Canadian context in order for sentences to “bolster, at the least, their capacity for denunciation, and, to some extent, deterrence”. He claims that, compared to the UK and Australia, Canadian judges are constrained in terrorism sentencing, which has led to wide discrepancies among analogous cases. In 2012, the Canadian landmark case of *R v. Khawaja* ruled that rehabilitation should not be minimized and could be a significant factor in terrorism sentencing (see for an analysis through the lens of Duff’s theory of punishment: Diab, 2013a). Zaia (2017) argues that it must be ensured that rehabilitation is imported as a mitigating circumstance, in accordance with *R v. Khawaja*. To judge otherwise, risks undermining deradicalization prospects (Zaia, 2017). Inextricably linked with the concept of rehabilitation is the assessment of dangerousness and the risk of recidivism. Whereas these risk assessments are normally made post-convictment, Berkell (2017) advocates for this assessment at sentencing level, as a factor affecting sentence type and sentence length. In other words, he argues for a utilitarian approach which is embedded in current trends “emphasizing data-driven analytics and evidence-based policies” (Berkell, 2017). Saif-Alden Wattad (2006) aligns with the finding that the notion of dangerousness should play a role at sentencing, in the sense that he argues that “terrorists are not more culpable or guilty of their offences because they are terrorists but they are more dangerous, and that is a factor

to be considered in sentencing”. Lee and Walker (2022) add that, in managing terrorism risk, the level of sentence execution and post-punishment must be given more attention.

Sentencing guidelines. Whilst the impact of sentencing guidelines is also mentioned in the next section, it is important to look into the findings of qualitative studies that have addressed the consequences of both US and UK sentencing guidelines in the context of terrorism (Brown, 2013; Floyd, 2021; Kelly, 2019; McLoughlin, 2010; Said, 2014; Sameer, 2016). These studies have often clarified or criticized the application of sentencing guidelines in one specific case – albeit with a more transcending conclusion. The US sentencing guidelines, also known as the “Terrorism Enhancement”, significantly increase the offense level and criminal history category in case of an act that involved, or was intended to promote, a federal crime of terrorism. Although initially mandatory, since the 2005 Supreme Court decision in the United States v. Booker, the guidelines are now considered advisory (Sameer, 2016). McLoughlin (2010) criticizes section 3A1.4 of the US Sentencing Guidelines, given that it “fails to provide for calibrating a defendant’s sentence to his or her conduct and characteristics”. As such, the guidelines undermine the principle that the punishment should be proportional to the crime (Sameer 2016; Floyd 2021). Brown (2013) adds that courts are often split on how to apply the terrorism enhancement, due to the conflict with goals of individualised sentencing (see also Floyd 2021). Burtis and Butler (2021) have studied what causes deviation from the guidelines and concluded that government manipulation of the guidelines and government’s view of a defendant play the biggest role in determining the final sentence length. Said (2014) pleads for the establishment of standards “to better help a court decide when a heightened punishment might be warranted, free from unsupported assumptions about the nature of terrorism or a particular defendant”. Floyd (2021), on the other hand, advocates for a revision of the Terrorism Enhancement to apply it more narrowly and exclude material

support offences of its scope. Kelly (2019) considers the UK Guidelines of the Sentencing Guidelines Council and its approach in assessing harm and criticizes the shift from an intent-based to a risk-based approach, since “both harm risked and harm intended are important to the seriousness of terrorism offences”. Lowe (2021), at his turn, delves specifically into “section 5 case trials” (which is the UK criminal law provision on preparation of terrorist acts) and studies how courts apply intent and aggravating factors within the UK Sentencing Guidelines.

Critiques. Post 9/11 prosecutorial strategies (in the US) are not free from criticism. The claims of successes are overshadowed by data-reliability and soft-sentence critiques (Chesney, 2007). On the one hand, there is disagreement on which cases should be labelled terrorism-related (namely the data-reliability critique; Chesney, 2007). On the other hand, the relatively short sentences override the claims of successes (described as the soft-sentence critique; Chesney, 2007). Chesney (2007) believes that case outcomes should always be examined on a per-offense basis (contrary to the broader label of terrorism-related), claiming that “offense-specific inquiries avoid the problems associated with the data-reliability critique, and permit more effective testing of the soft-sentence critique”.

Factors Influencing Sentencing Outcomes

The foregoing section has shed light on the judicial reasons behind sentencing terrorist offenders. The factors influencing those sentencing outcomes, as reported on in the literature, are highlighted below.

A. Terrorist versus non-terrorist offenders

Explained variance. While most studies solely focused on data of terrorist offenders, five early studies (all before 2010) focused on the sentence disparity between terrorist and non-terrorist offenders (Smith & Damphousse, 1996; Smith & Damphousse, 1998; Smith & Damphousse, 2002; Shields, Damphousse & Smith, 2006; Bradley-Engen, Damphousse & Smith, 2009). The main finding is that the ability to predict sentences of terrorists is substantially greater compared to sentences of non-terrorists, since political motive (and thus the official labelling of a ‘terrorist’) is found to be the most dominant explanatory variable for sentence length (Smith & Damphousse, 1996). A later study confirmed this finding in the sense that the explained variance for the terrorist sample was more than four times greater than the explained variance for the non-terrorist sample (Smith & Damphousse, 1998). In addition to political motive, the availability of additional variables on criminal history, crime severity and other crime-specific characteristics further increased the ability to predict sentencing outcomes (Smith & Damphousse, 1996; Smith & Damphousse, 1998).

The impact of US sentencing guidelines on this sentence disparity has been studied by Smith & Damphousse (2002) and Bradley-Engen, Damphousse & Smith (2009). The former found that while the sentence disparity in the pre-guideline era between terrorists and non-terrorists, and between terrorist group leaders and subordinates, was quite striking, this was no longer the case in the post-guideline era. The main finding of Bradley-Engen, Damphousse & Smith (2009) shows that terrorists continue to receive longer sentences than non-terrorists. However, the difference in average sentence length has declined significantly in the post-guideline era. As a result of the guidelines, the sentence disparity between terrorist and non-terrorist offenders has declined but the explained variance continues to be twice as much for terrorist offenders (Bradley-Engen, Damphousse & Smith, 2009).

Conviction at trial. Not only is the ability to predict sentences substantially greater in the case of a terrorist sample, according to a study of Shields, Damphousse & Smith (2006), terrorist defendants are more likely to go to trial and are twice as likely to be convicted at trial compared to a non-terrorist sample. Moreover, they found that in addition to the designation as a terrorist, only additional variables of age and the number of counts were found to be significant predictors of trial conviction.

B. Distinctions regarding types of terrorist offenders

The majority of studies do not make a comparison between a terrorist versus a non-terrorist sample but focus on specific aspects concerning the sentencing of exclusively terrorist offenders. This subsection will highlight the themes that focus on distinctions regarding types of terrorist offenders. While most studies include all types of terrorism within their research (and approach it as a variable of interest), there are some studies that focus on one type in particular (for example right-wing). Few studies focus on other dimensions, such as the difference between domestic and international terrorists or the behavioural variation across terrorist offenders.

Domestic versus international terrorists. A pre-9/11 study of Smith et al. (2002) showed, on the basis of a sample of 430 terrorist offenders between 1980 and 1998, that international terrorists (1) are more likely to have their crimes explicitly politicized; (2) are much less likely to plead guilty; and (3) when convicted, are punished more severely.

Behavioural variation. More recent studies have explored the diversity of behaviours that constitute involvement in terrorism (Harms, 2017; Horgan et al., 2018). While these studies provide extremely interesting insights into the qualities, complexity, and dynamic features of involvement in terrorism, this scoping review solely focuses on its

implications on the level of sentencing. Horgan et al. (2018), for example, distinguished between three types of behaviour (among individuals of the Global Jihadist Movement, convicted of terrorism-related offenses in the US), namely actors versus facilitators versus supporters, and found that actors received the most severe sentences, followed by facilitators. Supporters received the less severe sentences. Harms (2017), on the other hand, has compared conventional, violent terrorists to offenders of material support provisions (regardless of the ideological group). The sample of accomplices showed many differences compared to violent terrorists in the sense of ethnicity, education, and judicial treatment. Accomplices are more likely to accept a plea bargain, less likely to be convicted and less likely to be incarcerated when sentenced (Harms, 2017). Harms (2017) then categorized the sample of accomplices into four typologies, namely those who provide (1) finances (financers), (2) services (voluntary personnel), (3) physical materials (material providers), and (4) a combination of the aforementioned three. Of these typologies, financers receive a significantly shorter sentence. The study of Yon and Milton (2021) has also shown that whilst leaders do receive a longer sentence than non-leaders, “leaders were more likely to be treated with leniency, not severity, when it came to the seriousness of the charge they faced as well as whether law enforcement ultimately made a decision to continue the investigation itself”.

Right-wing terrorism. Another recent focus point in existing literature is on the disparate treatment between right-wing extremist violence and religiously inspired terrorism (Gruenewald et al., 2022; Nesbitt, 2021b; Rondon, 2018). Both in the United States (Rondon, 2018; Gruenewald et al., 2022) as well as in Canada (Nesbitt, 2021), authors argue that right-wing extremist violence is almost always treated as conventional crime or hate speech instead of through terrorism charges. Nesbitt (2021) argues that this finding results in a systemic bias, stigmatizing and punishing Al-Qaeda-inspired extremism more

than right-wing extremism. Gruenewald et al. (2022) clarify that federal prosecutors may avoid pursuing terrorism-related charges, since “they are historically less likely to result in convictions”. However, even though explicit terrorism charges are atypical in these cases, federal prosecutors often do mention the ideological or group affiliation during trial (Gruenewald et al., 2022). Gruenewald et al. (2022) have further delved into a US sample of right-wing extremists in a quantitative manner and found that “over 80% of violent RWEs are ultimately convicted in federal court, with slightly more than half (51.6%) of those convictions resulting from entered guilty pleas”. The average received sentence is 63 months, varying by conviction type and by both gender and age of the defendant.

C. Specific predictors of sentencing outcomes

Whereas the previous two subsections highlighted various themes related to a comparison between a terrorist versus a non-terrorist sample, or to distinctions within the terrorist sample, the current subsection will delve into a variety of specific factors that affect sentencing outcomes. These factors range from contextual to legal and offender-specific characteristics.

Major terrorist incidents. Multiple research articles have studied what the impact is of major terrorist incidents on sentencing outcomes (Amirault & Bouchard, 2017; Bradley-Engen et al., 2009; Damphousse & Shields, 2007; Yon & Milton, 2021). Contrary to common-sense beliefs, almost all studies confirm that when a terrorist offender is sanctioned in proximity to a major terrorist incident, the offender is punished less severely. This decrease in sentence length is the result of changing prosecutorial strategies in the aftermath of a terrorist event (see Yon & Milton, 2021). As Damphousse and Shields first proclaimed in 2007, there are three observations: (1) an increase of terrorism-

related indictments for relatively less serious offences; (2) single offenders are more likely to be charged for acts resulting in relatively short sentences; and (3) indictees are charged with fewer counts suggesting less complicated cases. Bradley-Engen et al. (2009) identified two underlying considerations: the discretion and incentive of prosecutors to negotiate sentence reductions for defendants in terrorist cases to increase the likelihood of conviction for terrorist leaders, and the prevention of future terrorist attacks. Amirault and Bouchard (2017) add one caveat: whilst contextual factors significantly affect sentencing outcomes, the observed decrease in sentence length is specific only to offenders motivated by Islamic extremist ideology. This finding indicates that sentencing outcomes are influenced not only by proximity to a major terrorist incident but also by changing threat perceptions. Amirault and Bouchard (2015) also found another relevant dimension of timing, namely that “offenders sanctioned in the later stages of a terrorist campaign are generally sentenced more severely than offenders adjudicated at the onset for similar crimes”.

Duration of case processing. Not only proximity to a major terrorist incident or the stages of a terrorist campaign but also other time-related factors may influence sentencing outcomes. Bradley-Engen et al. (2012) have investigated the previously unexplored effect of time “measured by the duration of case processing/time to conviction”. They found that time to conviction has a strong, direct effect on sentence length. The longer the duration of case processing, the more severe the sentence (more in particular, a 1% increase in time resulted in a 3% increase in sentence length). This “time penalty” demonstrates the importance of incorporating more procedural variables into sentencing studies. Moreover, time to conviction mediated the effects of mode of conviction, offender race, and to a lesser degree the effects of gender, criminal history and total number of counts.

Mode of conviction. Bradley-Engen et al. (2012) found that a trial conviction increases sentence length both directly (with over 70%) and indirectly by lengthening the duration of case processing. Moreover, their findings indicate that after crime severity, mode of conviction, time to conviction and their interaction present the strongest predictors of sentence length. Gruenewald et al. (2022) confirm this finding in relation to a sample of right-wing extremists, in which those defendants who opted to go to trial received longer sentences. This finding is quite unsurprising, given that “receiving shorter sentences is one of the major incentives for defendants to not pursue a trial by jury” (Gruenewald et al., 2022).

Type of evidence. The relationship between evidence, evidential value and sentence length is only covered by one study (Heinrich, Thornton, Morgan & Bouhana, 2012). Heinrich et al. found that the type of evidence differed in the context of IRA terrorist cases and Al-Qaeda inspired cases. Whereas in the former, ballistic evidence was the most characteristic type of evidence, digital and document evidence were used more frequently in case of the latter. Moreover, the study confirmed that there is a relationship between evidence and sentence length, in the sense that an increase in the total amount of evidence is related to an increase in sentence length. More elaborated models showed the relationship between evidence and sentence length per charge category (ranging from the most severe charge category 1 of murder to charge category 5 consisting of terrorist training). The authors concluded that “human biological evidence has the greatest evidential value, followed sequentially by ballistics, real and finally chemical evidence”. Although digital and document evidence were the most prevalent type of evidence, these are not as influential on sentence length.

Gender. Recent research has focused on the impact of gender on the sentencing outcomes of terrorist offenders (Alexander & Turkington, 2018; Galica, 2020; Jackson et al., 2021; Strømme, 2017; Weaver & Doty, 2021; Yon & Milton, 2021). Whereas the former two studies are the result of a literature review, the latter four represent an empirically founded analysis. These studies problematize gendered narratives and assumptions and the disparities in treatment in the judicial system. While these studies explore the role of gender in a terrorism context on multiple levels, the main findings relevant for this scoping review are that women are less likely to be arrested, (slightly) less likely to be convicted and receive more lenient sentences compared to male terrorist offenders (Alexander & Turkington, 2018; see also Weaver & Doty, 2021). Strømme (2017) adds that “the judgments and reporting around the issue area are often seeped in gendered commentary, presenting these women as misunderstood victims rather than motivated agents” (see also Galica, 2020). Jackson, Gruenewald and Ratcliff (2021) add that whilst their findings support that “females are often treated more leniently by the criminal justice system, they do not take into consideration the varying roles that women play in terrorist organizations and the degree to which they engage in political violence”. Contrary to these findings, Yon and Milton's analysis in 2021 did not provide strong evidence of leniency towards female offenders.

Citizenship, culture and religion. Amirault and colleagues (2015, 2016, 2017) note citizenship (with immigrants receiving longer sentences) and threat perceptions as significant factors. The work of Bielamowicz (2021) aligns with previous research, in the sense that both foreign citizenship and ‘otheredness’ increase the average sentence length, leading to the conclusion that “the process of prosecuting and combatting terrorism and political violence is inherently biased”. The most profound impact, according to this study, however, was the affiliation with a foreign terrorist organization. Rondon (2018)

confirms that “defendants in domestic terrorism cases face different classifications of the same crime depending on their religion or identity”. Frings et al. (2018) is the only study with an experimental design, in which it was found through an online survey that Muslims are more likely to be found guilty of terror crimes and to receive a more severe sentence than non-Muslims. Albeit not in a real-world situation, this study again showed a serious risk for systematic discriminatory bias. The research conducted by Yon and Milton in 2021 yields a nuanced perspective on the influence of race within the criminal justice system. Their findings affirm a tendency toward greater leniency in cases involving individuals of white racial backgrounds when it comes to the decision to pursue an investigation or to impose felony charges. However, it is noteworthy that this racial impact did not manifest itself to the same extent in the sentencing phase in terms of sentence length.

Other factors. Literature has shown that a diverse set of factors affect sentencing outcomes, including legal, extra-legal, and contextual factors. As evident from this results section, the majority of studies concentrate on specific aspects, while only a minority offer a comprehensive examination of factors significantly influencing case processing outcomes. Johnson (2012) presents such a general overview. His findings show that the number of counts is the primary predictor of a defendant to be prosecuted, tried, and convicted. Other significant factors are the type of offence (with racketeering as most likely offence to be convicted), number of co-defendants (the more co-defendants, the less likely to result in trial dispositions and criminal convictions), type of target, type of court (courts with a higher terrorism caseload are significantly less likely to dismiss terrorism-related cases), and type of terrorist group (with left-wing and international groups more likely to have their cases dismissed compared to right-wing groups) (Johnson, 2012). Yon and Milton (2021) found that “far left individuals tend to receive

higher sentences than jihadists”, but when it comes to the charges faced and decision to pursue an investigation individuals associated with jihadist ideology are treated more severely. McCann (2018) has also emphasized the relationship between sentencing outcomes and the ideology or political affiliation of panel members. Istiqomah (2020) also presents a general study on factors that contribute to sentencing of terrorism offences in Indonesia. Her findings suggest that variables related to harm to victims are the strongest predictors of sentence length. Being cooperative in the investigation process, good behaviour during trial and parenthood, on the other hand, appeared to be lenient factors in sentencing (Istiqomah, 2020). The level of cooperation was also studied in Burtis and Butler (2021), together with three other variables (namely guilty plea, level of regret and continued support for the ideology). As a result, four sentencing categories were developed representing a gradient scale, “with category 1 defendants being likely to get the minimum sentence length and category 4 defendants being likely to get the maximum”.

Calls For Future Research

The included studies have formulated numerous recommendations for future research. As Gruenewald et al. (2022) note, “there remains much to be learned about the factors shaping criminal justice decisions-making and official responses to those accused of committing ideologically motivated violence”. One of the most recurring recommendations relates to the differentiation *within* the sample of terrorist offenders. Discussions on ‘*the* terrorist offender’ are in need of further nuance, in terms of ideology (left-wing, right-wing, religious, ethno-nationalist or single issue terrorism) (see Smith & Damphousse, 1998), behavioural functions (Horgan et al., 2018; Yon & Milton, 2021 on differentiating between group leaders and plot leaders for example; also across genders: see Jackson et al., 2021) and specific charges (Chesney, 2007). In addition to

differentiation is an enlargement of the sample recommended (also over time, with an inclusion of more recent years) (Smith & Damphousse, 1998; Johnson, 2012), as well as the inclusion of additional variables. These variables relate to court and jurisdiction characteristics and judicial profiles (Smith & Damphousse, 1998; Johnson, 2012; Istiqomah, 2020; Bielamowicz, 2021), political climates and social contexts (Johnson, 2012; Istiqomah, 2020), specific characteristics on terrorist organizations (Johnson, 2012; Bielamowicz, 2021), pretrial and bail outcomes (Johnson, 2012), and completion of the offence (Amirault, et al., 2016). Moreover, whilst the thematic analysis has shown that some studies have compared a terrorist sample with a sample of non-terrorist offenders (especially before 2010), more recent studies have advocated for a similar approach. They have called for the inclusion of a “comparison group to explore how, or if, sentencing outcomes of terrorist offenders differ from general offending populations convicted of similar offences” (Amirault, et al., 2016; Amirault & Bouchard, 2017; see also Istiqomah, 2020). Gruenewald et al. (2022) confirm this recommendation by arguing that one must “comparatively examine how criminal justice responses to violent extremism compare to prosecutorial and judicial responses to parallel crimes that differ in motivational circumstances and offender affiliations”. Istiqomah (2020) also calls for more different types of qualitative methods, including interviews and ethnography (especially to study “how judicial profiles such as religious, social, economic as well as political profiles may affect terrorism trials”).

Discussion

This study has addressed the predominant research trends and methodologies used in the academic literature on sentencing terrorist offenders and has assessed how these aspects have evolved over time in relation to volume, nature, and characteristics. Whilst the majority of included studies report that existing research on the topic of sentencing

terrorist offenders is limited, this scoping review has shown that the body of literature has expanded since 2010. The growing volume of literature makes it pressing to present the key research trends of this complex topic, identify gaps in knowledge, and guide future research. Not only does this study advance the academic field of terrorism studies by providing a rigorous and transparent examination of the available literature, serving as a solid foundation for future research, but the thematic analysis can also inform both policy development and the practice of criminal justice in the context of terrorism.

The findings of this scoping review have the potential to inform policy and practice in two significant ways. First, findings related to the emergence of anticipatory prosecution, the use of trials as performative spaces and the tension between sentencing principles invite policymakers and practitioners to address these challenges and consider a more balanced approach to counter-terrorism legislation. Second, understanding the factors that influence sentencing outcomes can help legal practitioners and judges make more informed decisions, ensuring fairness and justice in the legal process. Specific recommendations are outlined in Table 2.

INSERT TABLE 2

In various respects, this scoping review contributes to academia and informs future research. Not only does this review address the “calls for future research” as outlined in existing literature, but it also demonstrates the need for further investigation and provides insights into how such research could be structured in terms of geographical scope, methodology, interdisciplinary collaboration, theoretical frameworks, dataset, relevant variables and statistical strategies. These recommendations are systematically discussed in this discussion section and are summarized in Table 3. Additionally, limitations of this scoping review are acknowledged here, offering opportunities for future research to address them.

INSERT TABLE 3

Main finding is that research within Europe is lacking, especially regarding civil law countries. The research area is well-established in the United States, and to a lesser extent in Canada, whilst studies in Europe are confined to a few examples in the United Kingdom, the Netherlands, Germany and Norway (n=9 in total). This underrepresentation is probably the result of the search strategy, since only English search terms were used. The author acknowledges that certain sources might be missed, especially national legal annotations to case law and reports or working papers written in the native language. Moreover, the use of single screening and coding in this study should be acknowledged as a limitation, as it may have introduced the potential for selection bias and the risk of missing relevant studies.

Regarding methodology, this scoping review has demonstrated that qualitative research gained significant weight since 2010. The use of mixed methods, however, is highly undervalued in existing research, even though combining legal and qualitative approaches with statistics presents stronger results, than either one of these methodologies on their own (Hall & Wright, 2008). Other methodological issues that surfaced in early research, are the use of open sources and the need for empirical databases. Today, open-source databases are developing at a fast pace, as is shown by the ‘Terrorism and Extremist Violence in the United States Database’ (TEVUS), which integrates four different datasets. These databases create opportunities for more elaborate, empirical research (Gruenewald et al., 2022), but are primarily US-based. Combined with the issue of US Sentencing Guidelines, this may also clarify to a certain extent the dominance of US studies within this scoping review. As a result, similar databases should be developed in other countries to enable data-driven research, with an emphasis on reliable, complete and recent data (Hardyns et al., 2020). Existing research primarily

relies on publicly available documents, including official data (such as court documents and legal records) as well as media reports. However, relying on publicly accessible documents, even when it concerns official data, can sometimes be problematic since not all countries publish judicial decisions in the same manner. In some countries, only highly significant cases are made public, which can distort the overall understanding of trends and practices related to sentencing. The author, therefore, acknowledges that the creation and maintenance of open-source databases come with some challenges, such as issues related to data accuracy, accessibility, standardization, and the need for continuous updates. Future research could delve deeper into the practical aspects of establishing open-source databases, including discussions on data collection methodologies, data verification processes, data sharing mechanisms, and strategies for addressing potential biases or gaps in the data. To advance data-driven research, fostering a multidisciplinary and collaborative approach is essential. To date, there has been limited multidisciplinary approach in the research, even though legal scholars could enhance the qualitative aspects of analysis, while social scholars contribute quantitative methods, creating a synergistic research environment.

An important caveat within this review is the fact that the findings must be placed within their national context. A large majority of the presented findings in the thematic analysis are set in the US context. As a result, not all themes are as relevant for other jurisdictions. Given the critiques in the literature on the definition of ‘terrorism-related’ prosecutions, one must be cautious with the interpretation of the results. As Berkell (2017) states, “to evaluate sentencing in terrorism prosecutions, one must first determine which charges and convictions to include in the analysis”. Moreover, the large timeframe of included studies – ranging from 1996 to 2022 – raises the question to what extent earlier findings are still applicable or relevant today. Terrorism is not a new phenomenon,

but has taken significantly different proportions in the last 25 years, in terms of threat levels, types, targets, etc. The changing nature of terrorist offences and offenders also has its impact on legislative developments and criminal justice responses, and vice versa. Given the importance of contextual and legislative developments, it is necessary to re-examine earlier findings in current times. There is definitely room for more in-depth research on specific subtopics, such as foreign fighters, right-wing extremism, the influence of gender, defence strategies, evidential issues, behavioural variation, and a comparison with a non-terrorist sample. Identifying patterns on the level of terrorism on the one hand, and patterns of judicial decision-making on the other hand, has a true value for criminal justice actors and policymakers in the field of counterterrorism.

Other limitations concern the research design and its implementation. A scoping review typically employs qualitative and narrative synthesis methods to achieve its objectives of mapping the literature and identifying research gaps. As a result, this research article is exploratory in nature and uses basic description which serves as a foundational step to understand the scope and context of the existing research landscape. Given the inclusion of various study types (heterogeneous studies), this type of research is not suited for more sophisticated quantitative methods but could form the basis for more systematic quantitative methods traditionally used in pivotal meta-analytical studies. As such, future research might examine the strength of relationships between certain factors and sentencing outcomes or identify patterns across studies, which might generate valuable insights and potentially uncover trends or patterns that may not be apparent through basic description alone. Moreover, the author acknowledges that research in this field is dynamic and that new studies are continuously emerging. Given

that the scoping review is based on available literature up to December 1, 2022, relevant studies may have been published since.⁶

In addition to methodological advancements, it is advisable for future research to focus specifically on exploring and testing existing theories (for example, conflict theory, consensus theory and/or structural-contextual theory) or developing new theoretical perspectives, since this scoping review has shown that relatively little attention is given to theoretical frameworks within this domain. Also, perspectives related to restorative justice, procedural justice, risk assessment models and labelling theory could be of added value in the research field.

This scoping review has shed light on the scientific literature on the sentencing of terrorist offenders, in terms of volume, nature and characteristics. With counter-terrorism research booming in the previous decades, this specific research domain has lagged behind. Nevertheless, serious efforts have been undertaken in the last 25 years, leading to a promising and growing research field. We can only encourage further research and recommend a mixed methods approach with larger databases, based on official, recent and offense-specific data, especially within European countries with a civil law tradition.

References

Included In Scoping Review (N = 59)

Alexander, A., & Turkington, R. (2018). Treatment of Terrorists: How Does Gender Affect Justice? *CTC Sentinel*, 11(8), 24-29.

⁶ See for example Gruenewald, J., Klein, B.R., Hayes, B.E., Parkin, W.S., & June, T. (2024). Examining Disparities in Case Dispositions and Sentencing Outcomes for Domestic Violent Extremists in the United States. *Crime & Delinquency*, 70(1), 126-156.

- Amirault, J., & Bouchard, M. (2015). A Group-Based Recidivist Sentencing Premium? The Role of Context and Cohort Effects in the Sentencing of Terrorist Offenders. *International Journal of Law, Crime and Justice*, 43(4), 512-534.
- Amirault, J., & Bouchard, M. (2017). Timing is Everything: The Role of Contextual and Terrorism-Specific Factors in the Sentencing Outcomes of Terrorist Offenders. *European Journal of Criminology*, 14(3), 269-289.
- Amirault, J., Bouchard, M., Farrell, G., & Andresen, M. A. (2016). Criminalizing Terrorism in Canada: Investigating the Sentencing Outcomes of Terrorist Offenders from 1963 to 2010. *Journal of Criminal Law and Criminology*, 106(4), 769-810.
- Berkell, K. (2017). Risk Reduction in Terrorism Cases: Sentencing and The Post-Conviction Environment. *Journal for Deradicalization*, 13, 276-341.
- Bielamowicz, I. (2021). The Impact of Foreign Affiliation and Citizenship on The Prosecution of Political Violence in the United States. In M. Loadenthal (Ed.), *Prosecuting Political Violence*. Routledge.
- Bradley-Engen, M. S., Damphousse, K. R., & Smith, B. L. (2009). Punishing Terrorists: A Re-Examination of U.S. Federal Sentencing in the Postguidelines Era. *International Criminal Justice Review*, 19(4), 433-455.
- Bradley-Engen, M. S., Engen, R. L., Shields, C., Damphousse, K. R., & Smith, B. L. (2012). The Time Penalty: Examining the Relationship Between Time to Conviction and Trial vs. Plea Disparities in Sentencing. *Justice Quarterly*, 29(6).
- Brown, G. D. (2012). Notes on a Terrorism Trial: Preventive Prosecution, Material Support and the Role of the Judge after United States v. Mehanna. *Harvard National Security Journal*, 4(1), 1-57.

- Brown, G. D. (2013). Punishing Terrorists: Congress, the Sentencing Commission, the Guidelines, and the Courts. *Cornell Journal of Law and Public Policy*, 23(3), 517-551.
- Burtis, M., & Butler, L. (2021). Friend or Foe? An Analysis of Factors Influencing Sentence Length in the Prosecution of Terrorism. In M. Loadenthal (Ed.), *Prosecuting Political Violence* (Pp. 24). Routledge.
- Chesney, R. M. (2007). Federal Prosecution of Terrorism-Related Offenses: Conviction and Sentencing Data in light of the Soft-Sentence and Data-Reliability Critiques. *Lewis & Clark Law Review*, 11(4), 851-902.
- Crowley, M. G. (2010). Tough on Terrorists: Challenging Traditional Sentencing Principles in Australia? *Federal Sentencing Reporter*, 22(4), 279-287.
- Damphousse, K. R., & Shields, C. (2007). The Morning After: Assessing the Effect of Major Terrorism Events on Prosecution Strategies and Outcomes. *Journal of Contemporary Criminal Justice*, 23(2), 174-194.
- De Goede, M., & De Graaf, B. (2013). Sentencing Risk: Temporality and Precaution in Terrorism Trials. *International Political Sociology*, 7, 313-331.
- De Graaf, B. (2016). The Hofstad Group on Trial: Sentencing the Terrorist Risk, 2005–2014. In B. De Graaf & A. P. Schmid (Eds.), *Terrorists on Trial. A Performative Perspective* (Pp. 371-418). Leiden University Press.
- De Graaf, B. (2019). Foreign Fighters on Trial: Sentencing Risk, 2013-2017. In *Radicalization in Belgium and the Netherlands* (Pp. 97-117).
- Diab, R. (2011). Sentencing for Terrorism Offences: A Comparative Review of Emerging Jurisprudence. *Canadian Criminal Law Review*, 15(3), 267-297.
- Diab, R. (2013a). R v. Khawaja and the Frought Question of Rehabilitation in Terrorism Sentencing. *Queen's LJ*(2), 587-619.

- Diab, R. (2013b). Sentencing of Terrorism Offences After 9/11: A Comparative Review of Early Case Law. In *Terrorism, Law and Democracy* (Pp. 347-381).
- Floyd, S. (2021). Irredeemably Violent and Undeterrable: How Flawed Assumptions Justify a Broad Application of the Terrorism Enhancement, Contradict Sentencing Policy, and Diminish U.S. National Security. *The Georgetown Law Journal Online*, 109, 142-172.
- Frings, D., Rice, K., & Albery, I. (2018). The Effects of Religion and Stereotype Content on Verdicts and Sentence Severity when Defending Terror Charges. *Applied Psychology in Criminal Justice*, 14(1), 37-54.
- Galica, N. (2020). Lifting the Veil: The Impact of Framing on Sentencing Disparities Between Males and Females Convicted of Terrorism since the Rise of ISIS. *UIC Law Review*, 53, 113-162.
- Gruenewald, J., Ratcliff, K., & Lucas, H. (2022). Criminal Justice Responses to Right-Wing Extremist (RWE) Violence in the United States. In B. Perry, J. Gruenewald, & R. Scrivens (Eds.), *Right-Wing Extremism in Canada and The United States* (Pp. 491-513).
- Harms, J. (2017). The War on Terror and Accomplices: An Exploratory Study of Individuals who Provide Material Support to Terrorists. *Security Journal*, 30(2), 417-436.
- Heinrich, D. P., Thornton, A. E., Morgan, R. M., & Bouhana, N. (2012). An Analysis of Forensic Evidence used in the Prosecution of Terrorism Cases in Britain between 1972 and 2008. *Policing*, 7(1), 96-108.
- Hemmingby, C., & Bjørge, T. (2016). The Trial and Sentencing. In C. Hemmingby & T. Bjørge (Eds.), *The Dynamics of a Terrorist Targeting Process: Anders B. Breivik and the 22 July Attacks in Norway* (Pp. 77-86). Palgrave Macmillan.

- Hoffmann, P. (2019). The “La Belle” Trial: The Sentencing of a Terrorist Bomber under the German Penal Code. *German Law Journal*, 6(3), 667-687.
- Horgan, J., Shortland, N., & Abbasciano, S. (2018). Towards a Typology of Terrorism Involvement: A Behavioral Differentiation of Violent Extremist Offenders. *Journal of Threat Assessment and Management*, 5(2), 84-102.
- Istiqomah, M. (2020). *Understanding Terrorism Sentencing Decisions in Indonesia: A Southern Criminology Approach* [University of New South Wales].
- Jackson, S. M., Ratcliff, K., & Gruenewald, J. (2021). Gender and Criminal Justice Responses to Terrorism in the United States. *Crime and Delinquency*.
- Johnson, B. D. (2012). Cross-Classified Multilevel Models: An Application to the Criminal Case Processing of Indicted Terrorists. *Journal of Quantitative Criminology*, 28(1), 163-189.
- Kelly, R. (2019). Sentencing Terrorism Offences: No Harm Intended. *Criminal Law Review*, 7.
- Lee, F., & Walker, C. (2022). The Sentencing of Precursor Terrorism Crimes in the United Kingdom. In C. Walker, M. L. Angli, & M. C. Melia (Eds.), *Precursor Crimes of Terrorism: The Criminalisation of Terrorism Risk in Comparative Perspective* (Pp. 316). Edward Elgar Publishing Limited.
- Lowe, D. (2021). Preparation of Terrorist Acts, Section 5 Terrorism Act 2006: Evidence Required and Sentencing Guidelines. *Expert Witness Journal*, 94-98.
- McCann, W. S. (2018). The Role of Sentencing in Criminal Appeals: An Exploratory Analysis of Terrorism Appeals (1988-2015). *University of La Verne Law Review*, 40(1), 1-30.
- McCarthy, A. (2004). Terrorism on Trial: The Trials of al Qaeda. *Case Western Reserve Journal of International Law*, 36, 513-522.

- McGarrity, N. (2013). 'Let the Punishment Match the Offence': Determining Sentences for Australian Terrorists. *International Journal for Crime and Justice*, 2(1), 1-15.
- McLoughlin, J. (2010). Deconstructing United States Sentencing Guidelines Section 3A1.4: Sentencing Failure in Cases of Financial Support for Foreign Terrorist Organizations. *Law & Inequality*, 28(1).
- Nesbitt, M. (2021a). Sentencing the Toronto 18: Lessons from Then, Lessons for Now. *Manitoba Law Journal*, 44(1), 370-397.
- Nesbitt, M. (2021b). Violent Crime, Hate Speech or Terrorism? How Canada Views and Prosecutes Far-Right Extremism (2001–2019). *Common Law World Review*, 50(1), 38-56.
- Nesbitt, M., & Hagg, D. (2020). An Empirical Study of Terrorism Prosecutions in Canada: Elucidating the Elements of the Offences. *Alberta Law Review*, 57(3), 595-648.
- Nesbitt, M., Oxoby, R., & Potier, M. (2019). *An Empirical and Qualitative Assessment of Terrorism Sentencing Decisions in Canada since 2001: Shifting Away from the Fundamental Principle and Towards Cognitive Biases* (IZA Discussion Papers).
- Pyne, A. (2011). The Mood and Temper of the Public: R v Lodhi and the Principles of Sentencing in the War on Terror. *Current Issues in Criminal Justice*, 23(2), 163-182.
- Roach, K. (2014). Be Careful What You Wish For? Terrorism Prosecutions in Post-9/11 Canada. *Queen's LJ*, 40(1), 99-140.
- Rondon, Y. (2018). Treatment of Domestic Terrorism Court Cases: Class and Mental Health in the Criminal System. *The American University Journal of Gender, Social Policy & the Law*, 26(2), 741-792.

- Said, W. E. (2014). Sentencing Terrorist Crimes. *Ohio State Law Journal*, 75(3), 477-528.
- Saif-Alden Wattad, M. (2006). Is Terrorism a Crime or an Aggravating Factor in Sentencing? *Journal of International Criminal Justice*, 4(5), 1017-1030.
- Sameer, A. (2016). Is History Repeating Itself? Sentencing Young American Muslims in the War on Terror. *The Yale Law Journal*, 126, 1520-1576.
- Scanlon, Z. (2014). Punishing Proximity: Sentencing Preparatory Terrorism in Australia and the United Kingdom. *Current Issues in Criminal Justice*, 25(3), 763-784.
- Shields, C., Damphousse, K. R., & Smith, B. L. (2006). Their Day in Court: Assessing Guilty Plea Rates among Terrorists. *Journal of Contemporary Criminal Justice*, 22(3), 261-276.
- Smith, B. L., & Damphousse, K. R. (1996). Punishing Political Offenders: The Effect of Political Motive on Federal Sentencing Decisions. *Criminology*, 34(3), 289-322.
- Smith, B. L., & Damphousse, K. R. (1998). Terrorism, Politics, and Punishment: A Test of Structural-Contextual Theory and the "Liberation Hypothesis". *Criminology*, 36(1), 67-92.
- Smith, B. L., & Damphousse, K. R. (2002). *American Terrorism Study: Patterns of Behavior, Investigation and Prosecution of American Terrorists*.
- Smith, B. L., Damphousse, K. R., Jackson, F., & Sellers, A. (2002). The Prosecution and Punishment of International Terrorists in Federal Courts: 1980-1998. *Criminology & Public Policy*, 1(3), 311-338.
- Strømme, E. E. J. (2017). *Jihadi Brides or Female Foreign Fighters? Women in Da'esh – From Recruitment to Sentencing* (GPS Policy Brief 01/2017).

- Weaver, M., & Doty, A. (2021). Gender, Jail, and Injustice: Gender Interaction Effects on Judicial Sentencing Rhetoric. In M. Loadenthal (Ed.), *Prosecuting Political Violence: Collaborative Research and Method*.
- Yon, R., & Milton, D. (2021). Examining the Outcome of Investigations and Prosecutions of Extremism in the United States. *Studies in Conflict & Terrorism*.
- Zaia, R. (2017). Mental Health Experts in Terrorism Cases: Reclaiming the Status of Rehabilitation as a Sentencing Principle. *Criminal Law Quarterly*, 64, 548.

Other References

- Arksey, H., & O'Malley, L. (2005). Scoping Studies: Towards a Methodological Framework. *International Journal of Social Research Methodology*, 19-32.
- Colquhoun, H., Levac, D., O'Brien, K., Straus, S., Tricco, A., Perrier, L., Kastner, M., & Moher, D. (2014). Scoping Reviews: Time for Clarity in Definition, Methods, and Reporting. *Journal of Clinical Epidemiology*, 67, 1291-1294.
- De Coensel, S. (2020). *Counter-Terrorism & Criminal Law. A Normative Legitimacy Test of Terrorism-Related Offences on Expression, Information and Movement*. Maklu.
- De Graaf, B., & Schmid, A. P. (2016). *Terrorists on Trial: A Performative Perspective*. Leiden University Press.
- Grant, M. J., & Booth, A. (2009). A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies. *Health Information and Libraries Journal*, 26, 91-108.
- Gruenewald, J., Klein, B.R., Hayes, B.E., Parkin, W.S., & June, T. (2024). Examining Disparities in Case Dispositions and Sentencing Outcomes for Domestic Violent Extremists in the United States. *Crime & Delinquency*, 70(1), 126-156.

- Gusenbauer, M., & Haddaway, N. (2020). Which Academic Search Systems are Suitable for Systematic Reviews or Meta-Analyses? Evaluating Retrieval Qualities of Google Scholar, Pubmed, and 26 other Resources. *Research Synthesis Methods*, 11(2), 181-217.
- Hall, M. A., & Wright, R. F. (2008). Systematic Content Analysis of Judicial Opinions. *California Law Review*, 96(1), 63-122.
- Hardyns, W., Snaphaan, T., & Dieussaert, J. (2020). Wat kunnen open bronnen ons leren over de Crime-Terror Nexus in België? *Panopticon*, 41(2), 160-180.
- Loadenthal, M. (2021). *Prosecuting Political Violence: Collaborative Research and Method*. Routledge.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *BMJ*.
- Munn, Z., Peters, M., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic Review or Scoping Review? Guidance for Authors when Choosing between a Systematic or Scoping Review Approach. *BMC Medical Research Methodology*, 18.
- Nesbitt, M., Roach, K., Hofmann, D., & Lee, K. (2021). Canadian Terror: Multi-Disciplinary Perspectives on the Toronto 18 Terrorism Trials. *Manitoba Law Journal*, 44(1).
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., et al. (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372(71).
- Peters, M. D. J., Godfrey, C., McInerney, P., Munn, Z., Tricco, A. C., & Khalil, H. (2020). Chapter 11: Scoping Reviews. In E. Aromataris & Z. Munn (Eds.), *JBIM Manual For Evidence Synthesis*. Joanna Briggs Institute.

Pham, M., Rajic, A., Greig, K., Sargeant, J., Papadopoulou, A., & McEwen, S. (2014).

A Scoping Review of Scoping Reviews: Advancing the Approach and Enhancing the Consistency. *Research Synthesis Methods*, 5, 371-385.

Van Der Woude, M. (2010). *Wetgeving in een Veiligheidscultuur: Totstandkoming van Antiterrorismewetgeving in Nederland gezien vanuit Maatschappelijke en (Rechts)Politieke Context*. Leiden University.

Figure 1. Study Flow Chart, as adapted from Page et al. (2021)

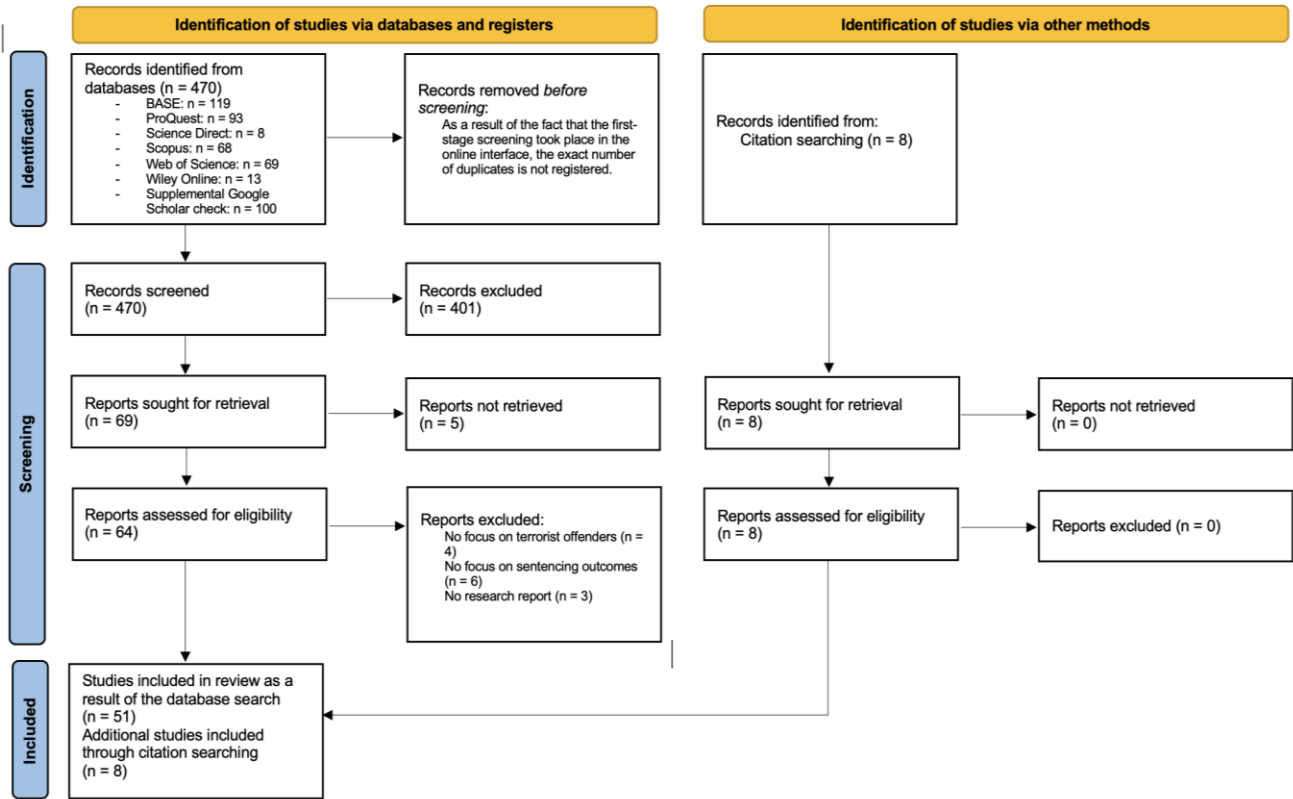


Figure 2. Number of publications over time

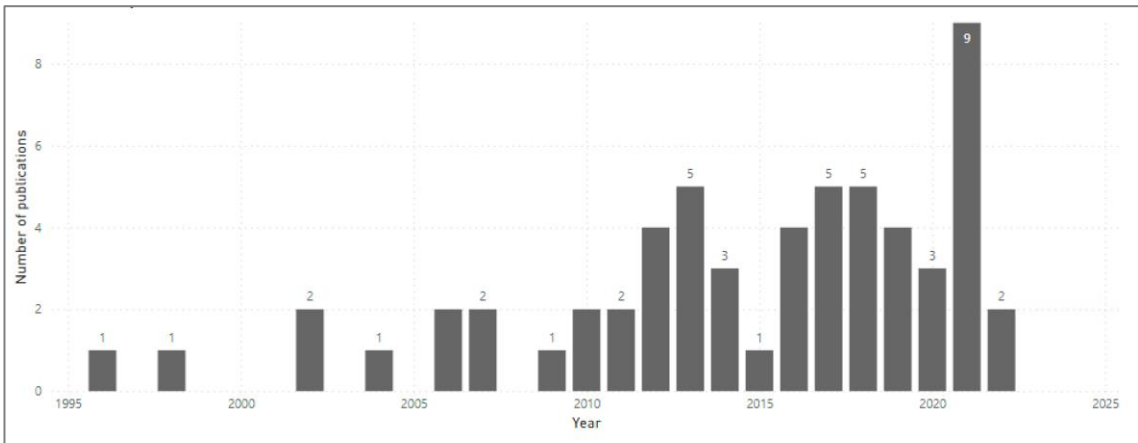


Figure 3. Geographic scope

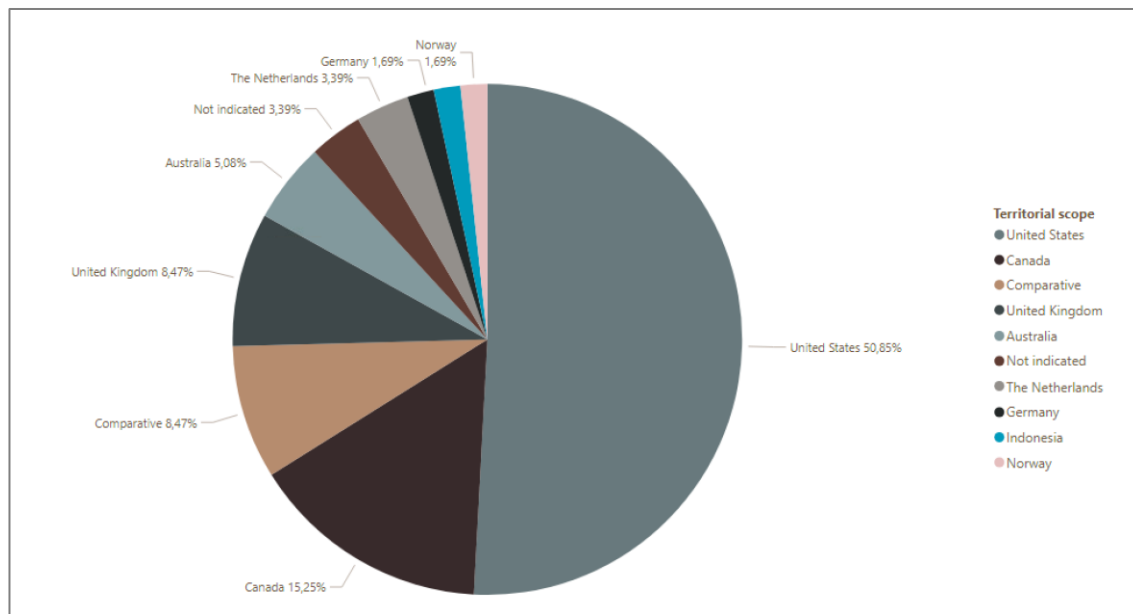


Table 1. Summary table quantitative & mixed methods studies

Author	Year	Territorial scope	Method	Timeframe	Sample size	Timespan
Smith & Damphousse	1996	United States	Inferential	Pre 9/11	<100	6-10y
Smith & Damphousse	1998	United States	Inferential	Pre 9/11	<100	6-10y
Smith & Damphousse	2002	United States	Inferential	Pre 9/11	>300	6-10y
Smith, Damphousse, Jackson & Sellers	2002	United States	Descriptive	Pre 9/11	>400	16-20y
Shield, Damphousse & Smith	2006	United States	Inferential	Pre 9/11	>400	6-10y
Chesney	2007	United States	Descriptive	Post 9/11	> 100	6-10y
Damphousse & Shields	2007	United States	Inferential	Pre & post 9/11	>500	16-20y
Bradley-Engen, Damphousse & Smith	2009	United States	Inferential	Pre 9/11	>100	6-10y
Bradley-Engen, Engen, Shield, Damphousse, Smith	2012	United States	Inferential	Pre & post 9/11	>400	21-25y
Heinrich, Thornton, Morgan & Bouhana	2012	United Kingdom	Inferential	Pre & post 9/11	>100	>30y
Johnson	2012	United States	Inferential	Pre 9/11	>500	21-25y
Amirault & Bouchard	2015	Canada	Inferential	Pre 9/11	>100	6-10y
Amirault, Bouchard, Farrell & Andresen	2016	Canada	Inferential	Pre & post 9/11	>100	>30y
Amirault & Bouchard	2017	United Kingdom	Inferential	Post 9/11	>100	6-10y
Harms	2017	United States	Inferential	Post 9/11	>700	6-10y
Frings, Rice & Albery	2018	Not indicated	Inferential	Post 9/11	>100	Not relevant
Horgan, Shortland & Abbasciano	2018	United States	Inferential	Pre & post 9/11	>100	16-20y
Nesbitt, Oxoby & Potier	2019	Canada	Mixed methods	Post 9/11	<100	16-20y
Galica	2020	United States	Mixed methods	Post 9/11	<100	0-5y
Istiqomah	2020	Indonesia	Mixed methods	Post 9/11	>100	11-15y
Bielamowicz	2021	United States	Descriptive	Post 9/11	>2000	16-20y
Burtis & Butler	2021	United States	Descriptive	Post 9/11	<100	16-20y
Jackson, Gruenewald & Ratcliff	2021	United States	Descriptive	Pre & post 9/11	>2000	>30y
Weaver & Doty	2021	United States	Inferential	Post 9/11	>100	16-20y
Yon & Milton	2021	United States	Inferential	Pre & post 9/11	>500	>30y
Gruenewald, Ratcliff & Lucas	2022	United States	Inferential	Pre & post 9/11	<100	26-30y

Table 2. Key findings & implications for policy and practice

Theme	Implications for policy and practice
Data sources	Improve data collection methods and systems within law enforcement agencies and relevant institutions to capture detailed information on individual terrorist acts and trials, in order to enhance data-driven research (and especially data analysis on a per-offence basis).
	Systematically disclose judgments and verdicts in terrorism cases to promote transparency and accountability. Ensure appropriate measures to safeguard the privacy and data protection of all individuals involved.
Legislation	Address overlap provisions and overbroad drafting in counter-terrorism legislation to ensure legal clarity, effectiveness, and respect for human rights.
	Avoid risk management as a central goal of new counter-terrorism legislation.
Terrorism trials	Prioritize the integrity of terrorism trials as a means to uphold the rule of law, protect human rights, and demonstrate the legitimacy of counter-terrorism measures.
Sentencing principles	Recognize the crucial role of rehabilitation in preventing recidivism and promoting the reintegration of individuals into society, rather than focusing on deterrence and denunciation.
	Enhance transparency and consistency in sentencing practices.
Sentencing guidelines	If sentencing guidelines are preferred, incorporate an intent-based approach, emphasizing the specific intent and actions of the offender rather than solely focusing on the perceived risk associated with the terrorist nature of the offense.
	Give preference to a case-by-case approach, taking into account individual circumstances, moral culpability, and the seriousness of the offense.
Factors influencing sentencing outcomes	Be aware of the dynamic nature of prosecutorial strategies in response to major terrorist incidents and evolving threat perceptions.
	Monitor the duration of case processing and adapt procedural frameworks to minimize the effects of “time penalties”.
	Recognize and address the potential for systemic bias in the application of terrorism charges, particularly when charges are disproportionately applied to cases with religious motivations.
	Ensure equal treatment of individuals, regardless of gender, citizenship status, ideology, or religion.

Table 3. Key findings & recommendations for scholars

Theme	Recommendations for scholars
Geographic scope	Prioritize research initiatives in countries with a civil law tradition to broaden the understanding of terrorism-related legal processes and sentencing practices beyond common law jurisdictions.
Methodology	Adopt mixed method approaches, combining qualitative and quantitative techniques, and draw from the benefits of triangulating data to enhance the robustness and validity of research findings.
	Undertake a meta-analysis and use sophisticated quantitative methods to examine the strength of relationships between certain factors and sentencing outcomes or identify patterns across studies.
Disciplines	Engage in interdisciplinary collaborations to explore multifaceted aspects of terrorism sentencing, leading to more mixed method approaches, comprehensive analyses, and innovative insights.
Theoretical framework	Advance theoretical frameworks, both by testing existing theories and constructing new conceptual models.
Data sources	Develop databases to enable data-driven research and delve into the practical aspects of establishing these open-source databases.
Coded variables & analysis	Analyze data on a per-offense basis to gain a more detailed understanding of sentencing outcomes.
	Think beyond sentence length as the sole dependent variable.
	Include more contextual variables.
	Explore distinctions within the terrorist sample to capture the diversity of cases and offenders.
	Re-examine earlier findings in light of the changing nature of terrorist offences, legislative developments and criminal justice responses.