

ROUTLEDGE CRITICAL SECURITY STUDIES SERIES

Researching Non-State Actors in International Security

Theory and Practice

Edited by
Andreas Kruck and
Andrea Schneiker



Researching Non-State Actors in International Security

This volume provides researchers and students with a discussion of a broad range of methods and their practical application to the study of non-state actors in international security.

All researchers face the same challenge, not only must they identify a suitable method for analysing their research question, they must also apply it. This volume prepares students and scholars for the key challenges they confront when using social-science methods in their own research. To bridge the gap between knowing methods and actually employing them, the book not only introduces a broad range of interpretive and explanatory methods, it also discusses their practical application. Contributors reflect on how they have used methods, or combinations of methods, such as narrative analysis, interviews, qualitative comparative analysis (QCA), case studies, experiments or participant observation in their own research on non-state actors in international security. Moreover, experts on the relevant methods discuss these applications as well as the merits and limitations of the various methods in use. Research on non-state actors in international security provides ample challenges and opportunities to probe different methodological approaches. It is thus particularly instructive for students and scholars seeking insights on how to best use particular methods for their research projects in International Relations (IR), security studies and neighbouring disciplines. It also offers an innovative laboratory for developing new research techniques and engaging in unconventional combinations of methods.

This book will be of great interest to students of non-state security actors such as private military and security companies, students of research methods, security studies and International Relations in general.

Andreas Kruck is Assistant Professor of Global Governance at Ludwig-Maximilians-University of Munich, Germany, and co-author of *International Organization* (2012).

Andrea Schneiker is Professor of Political Science at the University of Siegen, Germany, and author of *Humanitarian NGOs, (In)Security and Identity* (Routledge, 2015).

Routledge Critical Security Studies Series

Titles in this series include:

Reimagining War in the 21st Century

From Clausewitz to Network-centric Warfare

Manabrata Guha

The New Spatiality of Security

Operational Uncertainty and the US Military in Iraq

Caroline M. Croser

Human Security as Statecraft

Structural Conditions, Articulations and Unintended Consequences

Nik Hynek

US Domestic and International Regimes of Security

Pacifying the Globe, Securing the Homeland

Markus Kienscherf

Securitisation and the Iraq War

The Rules of Engagement in World Politics

Faye Donnelly

Ethics and Global Security

A Cosmopolitan Approach

Anthony Burke, Katrina Lee Koo and Matt McDonald

(In-)Security and the Production of International Relations

Securitisation and the Making of Foreign Affairs in Europe

Jonas Hagmann

Food Security Governance

Empowering Communities, Regulating Corporations

Nora McKeon

Biopolitics, Governmentality and Humanitarianism

‘Caring’ for the Population in Afghanistan and Belarus

Volha Piotukh

Ethical Security Studies

A New Research Agenda

Edited by Jonna Nyman and Anthony Burke

Securing Health

HIV and the Limits of Securitization

Suzanne Hindmarch

Researching Non-State Actors in International Security

Theory and Practice

Edited by Andreas Kruck and Andrea Schmeiker

Researching Non-State Actors in International Security

Theory and Practice

**Edited by Andreas Kruck and
Andrea Schneiker**

First published 2017
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2017 selection and editorial matter, Andreas Kruck and Andrea
Schneiker; individual chapters, the contributors

The right of the editor to be identified as the author of the editorial matter,
and of the authors for their individual chapters, has been asserted in
accordance with sections 77 and 78 of the Copyright, Designs and Patents
Act 1988.

The Open Access version of this book, available at www.taylorfrancis.com,
has been made available under a Creative Commons Attribution-Non
Commercial-No Derivatives 4.0 license.

Trademark notice: Product or corporate names may be trademarks or
registered trademarks, and are used only for identification and explanation
without intent to infringe.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book has been requested

ISBN: 978-1-138-94782-5 (hbk)

ISBN: 978-1-315-66983-0 (ebk)

DOI: 10.4324/9781315669830

Typeset in Times New Roman

Contents

<i>List of figures</i>	viii
<i>List of tables</i>	ix
<i>Notes on contributors</i>	x
<i>Preface and acknowledgements</i>	xvi
1 Introduction: researching non-state actors in international security – a multitude of challenges, a plurality of approaches	1
ANDREAS KRUCK AND ANDREA SCHNEIKER	
2 Conceptualizing political violence of non-state actors in international security research	14
ANDREAS ARMBORST	
PART I	
Interpreting texts	29
3 Rebels without a cause: narrative analysis as a method for research on rebel movements	31
ALEXANDER SPENCER	
4 PMSCs and Twitter: sentiment analysis as a tool for evaluating social media data	50
MAGNUS DAU AND MARLEN MARTIN	
5 Semi-structured interviews with non-state and security actors	65
ANJA MIHR	

6 Combining semi-structured interviews and document analysis in a study of private security expertise	81
JOAKIM BERNDTSSON	
7 Discussion chapter: comments on ‘interpreting texts’	96
JUTTA JOACHIM	
PART II	
Establishing causal claims	105
8 (Comparative) case studies: combining case study techniques for the causal analysis of security privatization	107
ANDREAS KRUCK	
9 Qualitative comparative analysis and the study of non-state actors	123
PATRICK A. MELLO	
10 Geospatial analyses of non-state actors in violent conflicts	143
ALEXANDER DE JUAN	
11 Discussion chapter: shadow boxing in Plato’s cave: assessing causal claims	159
BERTJAN VERBEEK	
PART III	
Doing fieldwork	169
12 An ethnographic approach to non-state security: participant observation among private security officers	171
TESSA DIPHOORN	
13 Using experimental methods in post-conflict countries to understand the effects of gender reforms in the Liberian National Police	187
SABRINA KARIM	
14 Empirical assessment of (policy) effectiveness: the role of business in zones of conflict	204
MELANIE CONI-ZIMMER AND KLAUS DIETER WOLF	

15 Discussion chapter: practising reflexivity in field research	222
--	------------

JACQUI TRUE

Conclusions and perspectives for researching non-state actors in international security	231
--	------------

16 From cookbooks to encyclopaedias in the making: methodological perspectives for research of non-state actors and processes	233
--	------------

ANNA LEANDER

<i>Index</i>	245
--------------	------------

Figures

2.1	Gradual model of political violence	23
4.1	Comparison of sentiment scores as percentages of all tweets in the data sets	57
9.1	Military participation and solution term	138
10.1	Shape and distribution of four different units of analysis used to analyse violence in Pakistan	147

Tables

4.1	Distribution of sentences per tweet in the data sets	55
9.1	INUS and SUIN causes	127
9.2	Example of a truth table	129
9.3	The fuzzy set ‘military participation’ in the Iraq War (outcome)	132
9.4	Fuzzy-set calibration	133
9.5	Truth table for ‘military participation’ in the Iraq War	136
9.6	Solution terms for ‘military participation’ in the Iraq War	137
14.1	Dimensions of effectiveness	206

Contributors

Andreas Armbrorst is head of the National Center for Crime Prevention in Germany, a think tank for evidence-based practices in criminal policy and counter-radicalization. In 2015 he was Marie Curie Fellow at the School of Law in Leeds, UK, where he investigated the long-term development of jihadi ideology. Previously he worked in the field of security and terrorism studies at the Max-Planck Institute for Foreign and International Criminal Law and the Center for Security and Society in Freiburg, Germany.

Joakim Berndtsson is a senior lecturer at the University of Gothenburg's School of Global Studies. His primary research interests are the privatization of war and security, but also include general security studies, the history of state formation, civil–military relations, the transformation of war and UN peacekeeping. Additionally, he is involved in a national research programme on Information Security, funded by the Swedish Civil Contingencies Agency, and works on a project on the public opinion of the Swedish Armed Forces. Recent work has been published in *Armed Forces and Society*, *Cambridge Review of International Affairs*, *Critical Military Studies* and the *Journal of Ocean Development and International Law*. Together with Christopher Kinsey, Joakim Berndtsson is also the editor of the *Routledge Research Companion to Outsourcing Security* (2016).

Melanie Coni-Zimmer is a senior researcher in the department on Private Actors in the Transnational Sphere at the Peace Research Institute Frankfurt (PRIF), Germany. Her research has focused on corporate social/security responsibility in conflict zones, global crime governance as well as corporations and natural resource governance. She has published several German and English monographs, articles and book chapters on the diffusion and effectiveness of corporate social responsibility in the field of security. She has also co-authored a policy report commissioned by the German Corporation for International Cooperation (GIZ, 2013) on *Friend or Foe? Developing Partnerships in Natural Resource Governance. A Global Stakeholder Analysis*.

Magnus Dau is a research associate at the University of Siegen, Germany. He works in the international collaborative research project 'Militarization 2.0:

Militarization's Social Media Footprint through a Gendered Lens', funded by the Swedish Research Council and directed by Susan Jackson (Stockholm), Nick Robinson (Leeds), Jutta Joachim (Nijmegen) and Andrea Schneiker (Siegen). Previously, he was a research associate and junior lecturer in Chinese politics at the University of Duisburg-Essen, Germany, and, before that, a research associate in comparative politics at the University of Marburg, Germany, where he also earned an MA in political science.

Alexander De Juan is a visiting professorial fellow for International Administration and Conflict Management at the University of Konstanz and a senior research fellow with the German Institute of Global and Area Studies (GIGA) in Hamburg. His current research interests include institutions in intrastate conflicts, the role of religion and ethnicity in civil war as well as external support to state-building in fragile states. Previously, he worked as a sector economist and project manager 'Peace and Security' for the German government-owned development bank KfW. Alexander De Juan's recent academic work has been published or is forthcoming in the *Journal of Peace Research*, *Journal of Conflict Resolution*, *Comparative Political Studies*, *Political Research Quarterly*, *British Journal of Political Science*, *Terrorism and Political Violence*, *Political Geography*, *Civil Wars* and *Conflict Management and Peace Science*.

Tessa Diphoorn is an anthropologist and assistant professor at the Department of Anthropology at Utrecht University. Previously she worked as a postdoctoral researcher at the Department of Human Geography, Planning and International Development Studies at the University of Amsterdam, the Netherlands, where she was part of a larger comparative project on security assemblages in Kingston, Nairobi and Jerusalem that was funded by a European Research Council (ERC) Starting Grant. Before that, she researched private security in Durban, South Africa, which resulted in a book – *Twilight Policing: Private Security and Violence in Urban South Africa* – with the University of California Press. In addition to that, she has published in various journals, such as *Policing and Society*, the *Journal of Contemporary Ethnography* and *Medical Anthropology*. In May 2017, she will start a new research project entitled 'Policing the police in Kenya: Analysing state authority from within'.

Jutta Joachim is professor of International Relations at Radboud Universiteit in Nijmegen, the Netherlands. Her research interests are: non-state actors in international relations, security governance, the Common Foreign and Security Policy of the EU, human rights and gender and international relations. She is the author of *Agenda Setting, the UN, and NGOs: Gender Violence and Reproductive Rights* (Georgetown University Press, 2007) and co-editor of *International Organizations and Implementation: Enforcers, Managers, Authorities and Transnational Activism in the UN and the EU: A Comparative Study* (both Routledge, 2008). Her articles have appeared in

International Studies Quarterly, the *German Journal for International Relations*, *Security Dialogue*, *Millennium*, *Cambridge Review of International Affairs*, *Critical Military Studies*, *Comparative European Politics* and the *Journal of European Public Policy*.

Sabrina Karim is an assistant professor in the Government Department at Cornell University in 2017. She is the co-author of *Equal Opportunity Peacekeeping*, published by Oxford University Press. Her research interests include gender reforms in the post-conflict security sector and in peacekeeping, the effect of security sector reform on peace and security, third party involvement in peace processes, and the relationship between conflict-related violence and post-conflict sexual violence. Much of her research has been in sub-Saharan Africa, where she has conducted field experiments, lab experiments, and surveys. Her studies have been published in the *International Organization*, *Journal of Peace Research*, *International Interactions* and *International Peacekeeping*, and her research has been funded by the National Science Foundation, the Folke Bernadotte Academy, the International Growth Centre, and the British Research Council.

Andreas Kruck is an assistant professor of Global Governance at Ludwig-Maximilians-University of Munich, Germany. He is the co-author (with Volker Rittberger and Bernhard Zangl) of *International Organization* (Palgrave, 2012) and the author of *Private Ratings, Public Regulations: Credit Rating Agencies and Global Financial Governance* (Palgrave, 2011). He has published on private actors, in particular private military and security companies and credit rating agencies, and international organizations in refereed international journals such as *The Annals of the American Academy of Political and Social Science*, *The Review of International Organizations*, *Journal of European Public Policy*, *Journal of International Relations and Development* and *Critical Studies on Security*.

Anna Leander is professor at the Copenhagen Business School, Denmark, and the Pontificia Universidade Católica do Rio de Janeiro, Brazil. Her research links theoretical debates about practice approaches in International Relations and empirical work on technological and commercial security practices. In her recent project Markets in the Making of Multilateral Military Interventions (4MI), funded by the Economic and Social Research Council and led together with Elke Krahmann, she analysed the implications of the interaction between international organizations and private military and security companies in multilateral military interventions. Her research has been published in leading journals such as *Review of International Studies*, *International Political Sociology*, *Security Dialogue*, *Review of International Political Economy*, *Journal of Peace Research* and *Millennium*. Anna Leander is currently completing a monograph on *Political Military Companies: Commercializing the Military and Military Politics* (under contract with Cambridge University Press). She is the editor of *The Commercialization of Security in*

Europe: Consequences for Peace and Reconciliation (Routledge, 2013) and co-editor (with Rita Abrahamsen) of the *Handbook on Private Military Security Studies* (Routledge, 2015).

Marlen Martin is a research associate at the Institute of Political Science at Leibniz University Hannover, Germany. She holds a degree in social sciences and works in the international collaborative research project ‘ Militarization 2.0: Militarization’s Social Media Footprint through a Gendered Lens’, funded by the Swedish Research Council and directed by Susan Jackson (Stockholm), Nick Robinson (Leeds), Jutta Joachim (Nijmegen) and Andrea Schneiker (Siegen). Marlen Martin’s research interests focus on the military, migration and national identity.

Patrick A. Mello is research and teaching associate at the Bavarian School of Public Policy at the Technical University of Munich, Germany. His research focuses on international security, the influence of domestic politics on foreign policy, and comparative and case study research methods (particularly fuzzy-set Qualitative Comparative Analysis). He is the author of *Democratic Participation in Armed Conflict* (Palgrave Macmillan, 2014) and has recently published in the *European Journal of International Relations*, the *Journal of International Relations and Development* and *West European Politics*.

Anja Mihr is founder and program director of the HUMBOLDT-VIADRINA Center on Governance through Human Rights in Berlin, Germany. She has held professorships at the Willy Brandt School of Public Policy at the University of Erfurt in Germany and at the Netherlands Institute of Human Rights (SIM), University of Utrecht, the Netherlands. She has been one of two principle investigators and research directors of the European ORA project on the Impact of Transitional Justice on Democratic Institution Building. Her work focuses on transitional justice, public policy, governance and human rights. She has published a number of books and articles on international human rights regimes, human rights education, transitional justice, democratization, the European human rights system and NGOs and has been co-editor of the *European Yearbook of Human Rights* as well as the *German Journal for Human Rights*.

Andrea Schneiker is junior professor of political science at the University of Siegen, Germany. She is the author of *Humanitarian NGOs, (In)Security and Identity: Epistemic Communities and Security Governance* (Routledge, 2015). Her research on state and non-state actors in international security has been published in numerous refereed international journals such as *Cambridge Review of International Affairs*, *Comparative European Politics*, *Disasters*, *International Studies Review*, *Millennium*, *Security Dialogue*, and *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*.

Alexander Spencer is associate professor at the Otto-von-Guericke University Magdeburg, Germany. His research centres on discourse-analytical methods

and violent non-state actors. His work has been published in journals such as *Foreign Policy Analysis*, *Security Dialogue*, *International Studies Perspectives* and the *Journal of International Relations and Development*. He has published a book with Palgrave entitled *The Tabloid Terrorist* in 2010, co-edited a volume on *Reconciliation after Terrorism* (Routledge, 2012) and a special issue in the *Journal of European Public Policy* on Fiascos in Foreign and Public Policy. His new monograph titled *Romantic Narratives in International Politics* was published with Manchester University Press in 2016.

Jacqui True is professor of Politics and International Relations and Australian Research Council Future Fellow in the School of Social Sciences at Monash University, Australia. She is also a Peace Research Institute Oslo (PRIO) Global Fellow. Her research is focused on understanding the political economy of post-conflict and the patterns of systemic sexual and gender-based violence in Asia Pacific conflict-affected countries. Recent publications include 'Reframing Conflict-Related Sexual Violence' in *Security Dialogue* with Sara E. Davies and *Scandalous Economics: The Politics of Gender and Financial Crises* (Oxford University Press, 2016) edited with Aida Hozic. With Brooke Ackery, she has written *Doing Feminist Research in the Political and Social Sciences* (Palgrave, 2010) which provides a methodological guide for researchers seeking to address the ethical dilemmas that inhere in the research process in social and political science.

Bertjan Verbeek is professor of International Relations at the Department of Political Science, Institute for Management Research at Radboud University, Nijmegen, the Netherlands. He is also a Visiting Researcher at Crismart, Swedish National Defense University, Stockholm, Sweden. His research interests include crisis decision making, the role of international organizations in world politics, and the impact of populism on foreign policy. His co-edited volumes include *International Organizations and Implementation: Enforcers, Managers, Authorities* (Routledge, 2008) and *Decision-making within International Organizations* (Routledge, 2004). Articles of his have been published in *Journal of Common Market Studies*, *European Journal of International Relations*, *Cooperation and Conflict*, *European Political Science Review* and *Acta Politica*.

Klaus Dieter Wolf is professor of International Relations at the Technical University Darmstadt and was formerly executive director of the Peace Research Institute Frankfurt (PRIF), Germany. He is the head of the PRIF research department on Private Actors in the Transnational Sphere. His recent research projects have focused on the role of transnational corporations in conflict zones, global crime governance, corporations and natural resource governance and the effectiveness and legitimacy of transnational governance. His work has appeared in *European Journal of International Relations*, *Global Constitutionalism*, *Cooperation and Conflict* and *Mershon International Studies Review*. He is co-author of *The Role of Business in Global Governance: Corporations as Norm-entrepreneurs*

(Palgrave Macmillan, 2010). He also co-edited *Corporate Security Responsibility? Private Governance Contributions to Peace and Security in Zones of Conflict* (Palgrave Macmillan, 2010), *The Transnational Governance of Violence and Crime: Non-State Actors in Security* (Palgrave Macmillan, 2013) and *Resistance and Change in World Politics* (Palgrave Macmillan, 2017).

Preface and acknowledgements

This volume brings together two major trends in research and teaching in the fields of Security Studies and International Relations (IR): a growing substantive interest in non-state actors and an increased emphasis on methods. Countless publications and university course modules of the past 20 years go beyond an exclusively state-centred understanding of international security. Non-state actors such as rebel groups, warlords, militias, terrorists, criminal groups, private military and security companies, local self-defence forces, business firms or non-governmental organizations (NGOs) regularly preoccupy practitioners, scholars and students of political science. These are not only increasingly popular but also particularly challenging objects (and subjects) of knowledge production for conceptual, practical and normative reasons (see our introduction to this volume in Chapter 1).

Precisely because non-state actors in international security pose a number of serious methodological problems, they constitute an exceptionally instructive field of research to probe a broad range of methodological approaches, including less traditional methods in IR such as narrative, ethnographic or geospatial analyses. This is why we asked an international group of researchers of non-state actors in international security to reflect on the practical application of a particular method (or combination of methods) in their own research. Moreover, we approached experts on these methods to provide a comparative discussion of the methods and their application, reflecting on the implications of choosing one method over another and discussing the relative merits and difficulties of the different approaches in use. The result is this book, which presents and discusses a broad and diverse selection of interpretive and explanatory methods ‘in use’, i.e. in studies of non-state security actors.

As it includes not only a presentation of selected methods and methodologies, but also engages in a discussion of their practical application as well as their merits and limitations, this volume fills a gap in the burgeoning literature and teaching on methods in IR and Security Studies. While students are generally taught different qualitative and quantitative methods and may also learn about how to choose a method for their own research, their textbooks and studies hardly ever prepare them for the challenges they confront when actually using them. Notwithstanding the boom of method books, classes and teaching materials

that has shaped IR and Security Studies, the leap from knowing methods to applying them is still a great one. Our ambition is to encourage, facilitate and help in training for this jump.

The idea for this volume grew out of several workshops of the research network on 'Privatisation of Security through Private Military and Security Companies', led by Andrea Schneiker and funded by the German Research Foundation (DFG). We gratefully acknowledge the financial support of the German Research Foundation. We particularly thank the members of the research network and the participants of workshops in Bonn, Munich and Siegen for their valuable comments on several chapters in this volume. We owe special thanks to Christiane Frantz, Anna Leander and Elke Krahmann who have supported this project with their intellectual ideas and practical advice. Our greatest thanks go to the authors of this volume, who have agreed to write pretty unconventional types of chapters and share their hands-on experiences when using methods in their own research. We hope a lot of students and scholars will find their open, transparent, and self-reflexive discussions of methods and their practical application useful.

Andreas Kruck and Andrea Schneiker
Munich/Siegen



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

1 Introduction

Researching non-state actors in international security – a multitude of challenges, a plurality of approaches

Andreas Kruck and Andrea Schneiker

From methodological textbook knowledge to practical research

Social science methods are tools that help us find new or better answers to important questions of our time. As is the case with most tools, knowing what a given method is good for and how it works in theory is one thing; being able to make the most productive use of it in research practice is quite another. To produce relevant and valid findings, researchers who conduct, for example, case studies must have a sound knowledge of case study methodology (see for example, George and Bennett 2005; Gerring 2007; Collier, Brady and Seawright 2010). The same holds true for post-positivist methods such as discourse and narrative analysis (Milliken 1999). But abstract methodological knowledge¹ is hardly sufficient; it must be put into practice to render the chosen method a powerful tool for generating theoretical or empirical insights.

All scholars of political science face the same challenge: not only must they identify a suitable method for analysing a research question; they must also apply it to a particular case or a number of selected cases. More often than not, it is at this point that things get interesting – but also tricky because there is a gap between knowing the methods and actually using them. In this edited volume, we hope to bridge this crucial gap by introducing methods of International Relations (IR) and political science ‘in use’. Fortunately, there are now plenty of excellent textbooks on methods – in IR in particular² and in the political and social sciences in general³ – that introduce, describe and explain various qualitative and/or quantitative methods, covering a broad range of explanatory and/or interpretivist approaches. However, most of these texts cannot sufficiently prepare us for the challenges that confront us when we actually apply these methods, because the authors usually focus on comprehensive descriptions and discussions of the different methods, rather than on their concrete application within specific research designs.

The chapters in this volume, by contrast, retrace and reflect on how researchers in IR and political science have used a broad range of methods to study non-state actors in international security in an effort to establish causal claims, analyse texts or do fieldwork. More specifically, the chapters address and offer guidance on three major questions.

- On the basis of which selection criteria can we choose a suitable method (or a combination of methods) to address a specific research question?
- How can we translate a method (or a combination of methods) into a research design?
- When we apply the chosen method(s) to a specific case, how can we deal with problems that arise during the research process?

Moreover, just as the man with a hammer tends to see a world full of nails, our knowledge and choice of a method – and, conversely, our ignorance and non-use of others – will condition what sort of knowledge we gain and what kind of insights we systematically neglect. Accordingly, the chapters in this volume highlight the implications of using a particular method for the results of the research, i.e. for the observations that are made and the interpretations or the explanations that are offered. The volume thereby introduces a broad range of methods and their applications within specific research contexts and projects, including both more traditional methods and newer ones (in IR, at least, and specifically with reference to non-state actors):

- Qualitative content analysis
- Qualitative interviews
- Narrative analysis
- Sentiment analysis
- Participant observation
- Causal case studies: covariational analysis, congruence analysis, process tracing
- Qualitative comparative analysis (QCA)
- Quantitative geospatial analysis
- Assessment of policy effectiveness
- Surveys
- (Lab-in-the-field) experiments.

Research on non-state actors in international security: proliferation of issues, multiple challenges and methodological ways forward

Not so long ago, the focus on non-state actors might have seemed odd to all but a few scholars of international security. After all, security studies were traditionally concerned with the causes of war and peace between states (Waltz 1979; Mearsheimer 2001; Russett and O’Neal 2001). While there had been earlier studies of non-state security actors (e.g. Eckstein 1964; Gurr 1970; Wardlaw 1982; Crenshaw 1983) and of the significance of transnational relations (Nye and Keohane 1971), it is especially since the end of the Cold War that a broad range of non-state actors have become highly relevant research objects. These actors include rebel groups, warlords, militias, terrorists, criminal groups, private military and security companies (PMSCs), local self-defence forces, business

firms and non-governmental organizations (NGOs), and they increasingly preoccupy both practitioners and scholars of political science and IR.⁴

Perhaps even more important than this proliferation of actors beyond the realm of states, their institutions and representatives, research issues and questions – as well as theoretical approaches to studying them – have also become more diverse (Buzan and Hansen 2009; Collins 2013; Williams 2013; Peoples and Vaughan-Williams 2015). Questions about what explains the behaviour of certain types of actors, e.g. terrorist attacks, or the interaction between non-state and state actors, e.g. (intra- as well as interstate) war and the peaceful resolution of conflicts, continue to be important concerns and are likely to keep researchers busy in the foreseeable future. But recent and contemporary research on non-state actors is not limited to questions about the causes and effects of particular actions and interactions. In current scholarship, a much broader set of novel research foci are being studied, and from a variety of theoretical perspectives, including explanatory and interpretive, positive and critical theories:

- How is it possible that certain security actors or policies are regarded as legitimate while others are considered illegitimate? (See Chapter 3 by Alexander Spencer, as well as Cutler 2010 and Krahmann 2012.)
- How are security threats and risks discursively constructed (see Buzan *et al.* 1998; Spencer 2010; Krahmann 2011), and what are the implications of different constructions?
- How do everyday micro-practices of non-state actors play out in conflict zones, and how do they shape the dynamics and outcomes of violent conflict? (See Chapter 12 by Tessa Diphooorn, as well as Abrahamsen and Williams 2011 and Hönke 2013.)
- When and how are non-state actors able to influence security policies, and what makes them powerful? (See Price 1998; Leander 2005a; de Jonge Oudraat and Hauffer 2008; Dembinski 2009.)
- How effective is security governance by non-state actors? (See Chapter 14 by Melanie Coni-Zimmer and Klaus Dieter Wolf, as well as Leander 2005b, Bryden and Caparini 2006, Deitelhoff and Wolf 2010 and Petersohn 2013.)

This list of still somewhat general research foci could be both extended and specified, but already reveals the proliferation of research objects and interests in the field of security studies, yielding new questions and novel insights. This is also reflected in the (more or less recent) establishment of academic journals such as *Civil Wars*, *Small Wars & Insurgencies*, *Critical Studies on Security*, *European Journal of International Security* and the *Journal of Global Security Studies*, which also cover non-traditional security actors and issues.

In important ways, non-state actors in international security are not only increasingly common but also particularly challenging objects and subjects of knowledge production, owing to both their non-state status and the peculiarities of the fields of security and security policy. With regard to the former, reliable and comparable data on non-state actors are often unavailable. Existing research

approaches and data sets are often still state-centric (see Chapter 9 by Patrick Mello on qualitative comparative analysis), for various reasons, one being the history of the discipline. In addition, non-state actors in armed conflicts often operate in inaccessible environments, so scholars can only rely to a limited extent on official documents, surveys or ready-made statistics. In Chapter 10, for example, Alexander De Juan shows how geospatial analyses that link geographical features with patterns of violence often work best for subnational units of analysis that are not congruent with state or administrative boundaries. The chapters by Joakim Berndtsson (Chapter 6) and Andreas Kruck (Chapter 8) show how researchers studying PMSCs face considerable challenges when trying to collect reliable data on non-state security actors with commercial interests; even if researchers are able to get official information about, say, the contracting of PMSCs by state agencies, the data might not be complete or comparable across different countries or both, thus posing methodological challenges.

In fact, the very distinction between non-state actors and state actors may be problematic. In international security, it is sometimes difficult to identify non-state actors and to distinguish them from state actors, because the former, such as NGOs, businesses (see Chapter 14 by Melanie Coni-Zimmer and Klaus Dieter Wolf) or even warlords, may provide 'public' functions and are therefore perceived as somehow 'public' actors (see Chapter 6 by Joakim Berndtsson), whereas state actors may follow 'private' (market) logics (see Chapter 8 by Andreas Kruck).

In addition, compared with state officials, non-state actors may be somewhat inexperienced objects of research. Whereas, at least, Western state bureaucracies generally have established rules for how to deal with requests from researchers, non-state actors, especially security actors, may not necessarily have appropriate protocols in place and hence may not know what information they can or should share with researchers. This raises questions about access to the field.⁵ Such inexperience and its implications are described in Anja Mihr's chapter on semi-structured interviews with representatives of NGOs (Chapter 5). Non-state actors in international security may also be particularly keen on increasing their (sometimes shaky) legitimacy and recognition, which shapes (and may distort) the way they present themselves and their actions in interviews (see Chapter 5 by Anja Mihr) or in public statements (see Chapter 4 by Magnus Dau and Marlen Martin and Chapter 6 by Joakim Berndtsson). In this sense, they may be prone to what one might call hidden agendas.

This issue also highlights the fact that scholars who study non-state actors should reflect on the politics of their methodological choices; that is, the performative effects of their research. To what extent do their selection of methods and their analysis contribute to the reification, normalization or questioning of certain assumptions about the nature and legitimacy of particular modes and actors in international security (see Chapter 3 by Alexander Spencer)? This is of specific relevance because non-state actors are often involved as both objects and active players in legitimization and delegitimation discourses about their appropriate role in international security (see especially Chapters 4, 5, 6 and 12).

These challenges of studying non-state actors are compounded and reinforced by the particular challenges of studying the field of security. The high stakes, as well as the resulting political sensitivity and the non-transparency or secrecy of security issues, further complicate access to reliable data and raise serious ethical questions (see for example Chapters 12 to 14 by Tessa Diphoorn, Sabrina Karim and Melanie Coni-Zimmer and Klaus Dieter Wolf respectively). In contrast to, for example, humanitarian NGOs that are frequently (whether correctly or not) perceived as being guided by altruistic motivations of alleviating the suffering of vulnerable populations, non-state actors in security, particularly armed ones, are not necessarily seen as the ‘good guys’. This is not least due to the fact that their actions might include violence; as a result, their legitimate actorhood may be challenged. Again, these perceptions highlight the need to be aware that our research ‘objects’ often have political agendas, as well as the potentially problematic performative effects of our own research in normative or political terms.

Although some or all of these methodological challenges may be particularly pertinent and difficult to overcome for non-state actors in security, we should not overestimate the latter’s specialness, let alone their uniqueness in methodological terms, not least because state and non-state actors are sometimes difficult to distinguish and we sometimes study both types within one research project. Most of the above-mentioned challenges also apply to studies of state actors and to at least some policy fields beyond security. Hence, they are not unique to non-state actors in international security; rather, it is precisely because non-state actors in international security pose a number of serious methodological challenges that ‘cracking this tough nut’ may help us learn a lot that is applicable to a broad range of research designs in IR and political science. Thus, research on non-state actors in international security offers plenty of challenges and opportunities to probe different methods and to reflect upon their best possible practical use. Such research is particularly instructive for students seeking insights and advice on how to best use specific methods in their research projects.

Given the above-named difficulties, this field of research also lends itself to the application of less traditional methods in International Relations, such as sociological and ethnographic approaches, in order to gain insights into, say, the experiences and self-understanding of private security actors, or to geospatial analyses in order to learn more about the formation and strategies of violent non-state actors. In this sense, the study of non-state actors in international security may be an innovative laboratory for developing new research techniques that may ultimately be fruitful for a broader range of research interests in IR (see, for example, the narrative analysis presented by Alexander Spencer in Chapter 3 and the sentiment analysis introduced by Magnus Dau and Marlen Martin in Chapter 4). Methods such as qualitative comparative analysis (presented by Patrick Mello in Chapter 9), which are commonly used in state-focused studies of IR and comparative politics, can offer new insights when applied to non-state actors.

Given these expanded research opportunities and challenges, it becomes even more imperative to select a suitable method for pursuing a particular research goal and to apply it in the most productive way possible. This approach points to

one of the major lessons – if not *the* major lesson – to be learned from the chapters compiled in this volume: the research interest or question, i.e. the analytical problem, should determine the choice of method or methods – not the other way around. This may sound trivial, but time and again we seem tempted to conduct primarily methods-driven research rather than starting out with an important research question and then selecting the most productive method(s) to answer it.

From this plea for a problem-driven choice of methods follows a second key tenet reflected in this volume: as the study of non-state actors in international security involves a multitude of legitimate and productive research objectives, and as different research objectives require different methods, a plurality of methods – quantitative and qualitative, explanatory and interpretivist, positivist ‘problem-solving’ and ‘critical’ (see Cox 1981) – can, should and does inform research on non-state actors in international security. Even though a single volume can hardly do justice to the existing range of practised and promising methods for examining non-state actors in international security, we have chosen as broad and diverse a selection as possible.

The volume as a whole is not confined to any normative outlook or epistemological predisposition, nor does it favour a particular theoretical or methodological approach – it provides theory-based but also very practical guidance on the selection and application of a method embedded in a certain research design suited to answer a specific research question. In doing so, the volume goes beyond traditional dichotomies such as quantitative versus qualitative or positivist versus post-positivist methods. This does not mean that it is blind to important epistemological and sometimes also political differences implied by the selection and use of different methods. Quite the contrary, it is all about making conscious and informed choices: what kind of method will help me gain the particular kind of knowledge I seek, and which types of knowledge are necessarily and systematically overlooked or ignored by my choice of a specific method?

Such selectivity in the knowledge-producing capacity of a specific method can, at least to a certain extent, be mitigated by mixed-methods designs, which combine several methods of data generation and data analysis. Thus, a call for methods triangulation is a third major lesson provided by this volume, which features several chapters that combine two or more methods, acknowledging that the daily needs of researchers are often more complex than single-method approaches would suggest and that, in times of widely cherished methods triangulation, many types of analyses involve more than one method. For example, Chapter 4 by Magnus Dau and Marlen Martin on PMSCs and their (self-)presentation on Twitter combines sentiment analysis with a more qualitative framing analysis, whereas Chapter 6 by Joakim Berndtsson shows how interviews and qualitative content analysis can complement each other as data-generating tools in a study of state contracting with PMSCs. In Melanie Coni-Zimmer and Klaus Dieter Wolf’s research (Chapter 14), evidence on the effectiveness of corporate engagement in zones of conflict was gathered through a combination of semi-structured interviews with representatives of businesses and a diverse set of stakeholders, content analysis of the publicly

available reports of corporations and stakeholders and observations during field research in conflict zones.

Methods can be combined sequentially, as in the chapter on different case-study approaches by Andreas Kruck (Chapter 8); alternatively, they may be nested, as revealed through Tessa Diphoorn's research design, in which interviews were integrated into the broader practices of participant observation (see Chapter 12). Methods may be combined either on more or less equal footing (see Chapter 13 on surveys and experiments by Sabrina Karim and Chapter 8 by Andreas Kruck on case study approaches) or with one method being supplementary to the other, main approach (see Chapter 12 by Tessa Diphoorn on participant observation complemented with interviews). At any rate, as will be shown in the chapters, combining methods requires careful attention to (potentially) different data requirements and epistemological outlooks and, almost certainly, additional work. But, as the research examples presented in this volume show, these efforts may often pay off in terms of generating broader and/or deeper insights than a single-method design would allow.

Choosing and using methods: some questions as guidelines

Based on our discussion so far, and as will be emphasized in the chapters that follow, it should be apparent that there is no optimal and generally applicable formula for choosing and using the 'right' method. Nonetheless, the research experiences described in this volume do suggest a number of thematically clustered questions that may be regarded as a set of guidelines for making conscious and informed methodological choices both at the beginning of a research project and 'along the way'.

Research interest and theoretical approach

- Does a specific method or combination of methods allow for the analysis of the actors, structures or processes in which we are interested?
- Does it fit with the theoretical approach of our analysis?
- Is the chosen method appropriate for the type of empirical data we have at our disposal for analysing our research question?
- Is the method suitable for generating the epistemological type of knowledge in which we are interested?

Skills and resources

- Which skills on the part of the researcher are required by a specific method or set of methods?
- What technical equipment (e.g. hardware or software) is necessary? Do we have the technical skills to use it?
- How much time do we have for completing the research? How long will it take to apply a specific method of data generation and analysis?

Implications of chosen method(s)

- How will the outcome of the analysis in our cases be influenced by the chosen method(s)?
- What information will we get by using a certain method, and what information will we not get?
- What are the political and/or ethical implications of our analysis? Does our research have problematic performative effects?

Single-method vs. multi-methods designs

- What is the added value of combining several methods? Will it be enough to justify the extra work involved by the use of most multi-methods designs?
- Are the methods compatible and/or complementary in terms of data requirements, analytical insights and basic epistemological outlooks?

Coping strategies

- Have we clarified and justified all major choices and changes made to our research design that proved to be necessary during the course of the research process?
- Do we have a methodological second-best or fall-back option in case a certain method and research strategy do not generate data that will allow meaningful inferences (for analytical or practical reasons)?
- Have we made sure that competent and sympathetic supervisors and/or peers will provide feedback and advice regarding the progress of the research project?

This is not a definitive and comprehensive checklist. Other, additional issues may be relevant to individual research projects. However, most research projects will benefit if researchers consciously ask themselves and address these questions. Nevertheless, no book on methods can make up for one crucial aspect: training. None of us who seeks to apply a particular method or combination of methods for the first time should expect to be an expert on these methods from the beginning; rather, the more often one applies a method, the better one is likely to get at doing so. Hence, training ourselves in the practical use of methods is both a necessary and a rewarding endeavour.

Scope and structure of this book

In order to provide some ‘empirical overlap’ (Klotz 2008: 2) and to permit a meaningful comparison of the various methods ‘in use’, this book focuses on one common field of research, namely non-state actors in international security. Instead of treating different methodological approaches as individual and independent of one another, this volume discusses different methods by offering

examples of research that involve a wide range of non-state actors in international security.

Authors who themselves have conducted research on non-state actors each present a particular method or set of methods; that is, its selection and application. These contributions cover methods that focus on the interpretation of text (Chapters 3 through 6); seek to establish a causal claim (Chapters 8 through 10); and involve some fieldwork (Chapters 12 through 14). They include small-*n*, medium-*n*, and large-*n* approaches, as well as numerous instances where several kinds of methods have been combined in a multi-methods research design.

Before different methods and their applications are discussed, in Chapter 2, Andreas Armbrorst focuses on terrorists and terrorist acts to illustrate the conceptualization of non-state actors in international security and to discuss the difficulties in defining 'terrorism' and 'terrorists' and the profound and lasting consequences for the overall research project of conceptual decisions made at the outset.

The chapters in Part I, 'Interpreting texts', feature narrative analysis (Chapter 3), sentiment and frame analysis (Chapter 4), interviews (Chapters 5 and 6) and qualitative content analysis (Chapter 6). In Chapter 3, Alexander Spencer conducts a narrative analysis of the discourse in the British media on the rebellion and revolution in Libya in 2011 and shows that this type of analysis helps us understand how certain definitions of non-state security actors become dominant whereas others are marginalized. While in this case the text corpus that was analysed could be reduced to a reasonable size, allowing for a qualitative analysis, Magnus Dau and Marlen Martin (see Chapter 4), in contrast, had to cope with big data: 264,852 tweets on the operations of the PMSC Academi Elite. In order to interpret this vast amount of data, they first carried out a sentiment analysis that allowed them to identify the most relevant tweets, which were then further analysed by means of a qualitative content analysis. Another method often used to generate textual data are qualitative interviews. In Chapter 5, Anja Mihr draws on her research on non-state actors in the context of transitional justice across different countries such as Sierra Leone, Brazil, Hungary, Japan, Spain and Germany to discuss how to make the most of semi-structured interviews. In Chapter 6, Joakim Berndtsson introduces a study involving a contract between the Swedish Foreign Ministry and a PMSC to demonstrate how semi-structured interviews with state officials and representatives of the private security industry can complement thematic (content) analysis in order to obtain information that cannot be found in the official documents published by these actors.

Part II, 'Establishing causal claims', presents case study techniques (Chapter 8), qualitative comparative analysis (QCA) (Chapter 9) and geospatial analysis as an innovative form of quantitative analysis (Chapter 10). In Chapter 8, Andreas Kruck discusses how to combine the case study techniques of covariational analysis, congruence analysis and causal process tracing to explain security privatization in the United States, the United Kingdom, France and Germany. In Chapter 9, Patrick Mello demonstrates how a medium-sized sample of cases can fruitfully be analysed based on qualitative comparative analysis.

Although this method has been used most often to explain state actions and interactions, he discusses how it can also contribute to the study of non-state actors in international security. In Chapter 10, Alexander De Juan presents geospatial analysis as a promising quantitative tool to establish correlations among spatial variables, socioeconomic factors and violence even in highly challenging research settings, in this case in his analysis of militias in Darfur.

Part III, 'Doing fieldwork', introduces research approaches that involve varying amounts of fieldwork. In Chapter 12, Tessa Diphooorn discusses guidelines for using participant observation based on her studies of private security officers in South Africa, Kenya, Jamaica and Israel. In Chapter 13, Sabrina Karim elaborates on how to combine different kinds of experiments and surveys to analyse the effects of public security policies on the Liberian population. Chapter 14 by Melanie Coni-Zimmer and Klaus Dieter Wolf presents empirical assessments of the effectiveness of business company policies in conflict zones such as Nigeria, Rwanda and the Democratic Republic of the Congo based not only on observation during field research but also on semi-structured interviews and content analysis of publicly available reports.

This book not only includes a presentation of selected methods but also engages in a discussion of their practical application and their merits and limitations as compared with other approaches. In addition to the methods chapters, each section ends with a discussion chapter in which expert authors provide a comparative reflection on the methods presented in the relevant section and their application: Jutta Joachim discusses the methods applied in the section on 'Interpreting texts' (Chapter 7), Bertjan Verbeek reflects on the different approaches presented in the section on 'Establishing causal claims' (Chapter 11) and Jacqui True discusses the different ways of 'Doing fieldwork' (Chapter 15). The authors of these discussion chapters reflect on the implications of choosing one method over another and discuss the relative merits and difficulties of the different approaches in use. The volume concludes with a chapter by Anna Leander (Chapter 16) on methodological perspectives for research on non-state actors, calling for 'an approach to method that is well informed, but also open, imaginative and ready to make unconventional methodological moves and combinations'.

Notes

- 1 Whereas 'methods' denotes the tools or means chosen to collect and/or analyse data, 'methodology' refers to the theory and systematic analysis of the body of methods used in a particular (sub-)discipline.
- 2 See Sprinz and Wolinsky-Nahmias (2005), Ackerley, Stern and True (2006), Klotz and Prakash (2008), Salter and Mutlu (2013), Shepherd (2013), Stump and Dixit (2013), Aradau *et al.* (2015) and Lamont (2015), among others.
- 3 See King, Keohane and Verba (1994), Van Evera (1997), George and Bennett (2005), Yanow and Schwartz-Shea (2006), Gerring (2007, 2012), Box-Steffensmeier, Brady and Collier (2008), Hancké (2009), Kellstedt and Whitten (2009), Rihoux and Ragin (2009), Schatz (2009), Brady and Collier (2010), Blatter and Haverland (2012), Schneider and Wagemann (2012) and Bennett and Checkel (2015), among others.

- 4 Relevant recent articles and books on these and other non-state security actors abound. For an overview with additional references, see, among others, the contributions in Dunn Cavelty and Mauer (2009), Burgess (2010), Schmid (2013), Newman and DeRouen (2014) and Abrahamsen and Leander (2016).
- 5 Nevertheless, non-state actors may offer insights (e.g. into their motivations and justifications for action) that bureaucratic actors may be unwilling to disclose because the latter are bound to conceptions of 'national interest' and 'raison d'état' (see Chapter 5 by Anja Mihr).

References

- Abrahamsen, Rita, and Anna Leander, eds. 2016. *Routledge Handbook of Private Security Studies*. Abingdon: Routledge.
- Abrahamsen, Rita, and Michael C. Williams. 2011. *Security beyond the State: Private Security in International Politics*. Cambridge: Cambridge University Press.
- Ackerly, Brooke A., Maria Stern, and Jacqui True, eds. 2006. *Feminist Methodologies for International Relations*. Cambridge: Cambridge University Press.
- Aradau, Claudia, Jef Huysmans, Andrew Neal, and Nadine Voelkner, eds. 2015. *Critical Security Methods: New Frameworks for Analysis*. Abingdon: Routledge.
- Bennett, Andrew, and Jeffrey T. Checkel, eds. 2015. *Process Tracing: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press.
- Blatter, Joachim, and Markus Haverland. 2012. *Designing Case Studies: Explanatory Approaches in Small-N Research*. Houndmills: Palgrave Macmillan.
- Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier, eds. 2008. *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.
- Brady, Henry E., and David Collier, eds. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. 2nd edn. Lanham, MD: Rowman & Littlefield.
- Bryden, Alan, and Marina Caparini, eds. 2006. *Private Actors and Security Governance*. Vienna: Lit.
- Burgess, J. Peter, ed. 2010. *The Routledge Handbook of New Security Studies*. Abingdon: Routledge.
- Buzan, Barry, and Lene Hansen. 2009. *The Evolution of International Security Studies*. Cambridge: Cambridge University Press.
- Buzan, Barry, Ole Wæver, and Jaap de Wilde. 1998. *Security: A New Framework for Analysis*. Boulder: Lynne Rienner.
- Collier, David, Henry E. Brady, and Jason Seawright. 2010. 'Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology', pp. 161–199 in *Rethinking Social Inquiry*, H.E. Brady and D. Collier, eds. Lanham, MD: Rowman & Littlefield.
- Collins, Alan. 2013. *Contemporary Security Studies*. 3rd edn. Oxford: Oxford University Press.
- Cox, Robert W. 1981. 'Social Forces, States and World Orders: Beyond International Relations Theory'. *Millennium: Journal of International Studies* 10(2): 126–155.
- Crenshaw, Martha, ed. 1983. *Terrorism, Legitimacy, and Power: The Consequences of Political Violence*. Middletown: Wesleyan University Press.
- Cutler, A. Claire. 2010. 'The Legitimacy of Private Transnational Governance: Experts and the Transnational Market for Force'. *Socio-Economic Review* 8(1): 157–185.
- Deitelhoff, Nicole, and Klaus Dieter Wolf, eds. 2010. *Corporate Security Responsibility? Corporate Governance Contributions to Peace and Security in Zones of Conflict*. Houndmills: Palgrave Macmillan.

- de Jonge Oudraat, Chantal, and Virginia Haufler. 2008. *Global Governance and the Role of NGOs in International Peace and Security*. AICGS Policy Report No. 33. Washington, DC: American Institute for Contemporary German Studies.
- Dembinski, Matthias. 2009. 'NGOs and Security: The Case of the European Union', pp. 154–168 in *Transnational Activism in the UN and the EU: A Comparative Study*, J. Joachim and B. Locher, eds. Abingdon: Routledge.
- Dunn Cavelty, Myriam, and Victor Mauer, eds. 2009. *The Routledge Handbook of Security Studies*. Abingdon: Routledge.
- Eckstein, Harry, ed. 1964. *Internal War: Problem and Approaches*. New York: Free Press.
- George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Gerring, John. 2007. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.
- Gerring, John. 2012. *Social Science Methodology: A Unified Framework*. 2nd edn. Cambridge: Cambridge University Press.
- Gurr, Ted R. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
- Hancké, Bob. 2009. *Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences*. Oxford: Oxford University Press.
- Hönke, Jana. 2013. *Transnational Companies and Security Governance: Hybrid Practices in a Postcolonial World*. Abingdon: Routledge.
- Kellstedt, Paul M., and Guy D. Whitten. 2009. *The Fundamentals of Political Science Research*. Cambridge: Cambridge University Press.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Klotz, Audie. 2008. 'Introduction', pp. 1–7 in *Qualitative Methods in International Relations: A Pluralist Guide*, A. Klotz and D. Prakash, eds. Houndmills: Palgrave Macmillan.
- Klotz, Audie, and Deepa Prakash, eds. 2008. *Qualitative Methods in International Relations: A Pluralist Guide*. Houndmills: Palgrave Macmillan.
- Krahmann, Elke. 2011. 'Beck and Beyond: Selling Security in the World Risk Society'. *Review of International Studies* 37(1): 349–372.
- Krahmann, Elke. 2012. 'From "Mercenaries" to "Private Security Contractors": The (Re)Construction of Armed Security Providers in International Legal Discourses'. *Millennium: Journal of International Studies* 40(2): 343–363.
- Lamont, Christopher. 2015. *Research Methods in International Relations*. London: Sage.
- Leander, Anna. 2005a. 'The Power to Construct International Security: On the Significance of Private Military Companies'. *Millennium: Journal of International Studies* 33(3): 803–826.
- Leander, Anna. 2005b. 'The Market for Force and Public Security: The Destabilizing Consequences of Private Military Companies'. *Journal of Peace Research* 42(5): 605–622.
- Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York: Norton.
- Milliken, Jennifer. 1999. 'The Study of Discourse in International Relations: A Critique of Research and Methods'. *European Journal of International Relations* 5(2): 225–254.
- Newman, Edward, and Karl DeRouen, Jr., eds. 2014. *Routledge Handbook of Civil Wars*. Abingdon: Routledge.
- Nye, Joseph S., and Robert O. Keohane. 1971. 'Transnational Relations and World Politics: An Introduction'. *International Organization* 25(3): 329–349.

- Peoples, Columba, and Nick Vaughan-Williams. 2015. *Critical Security Studies: An Introduction*. Abingdon: Routledge.
- Petersohn, Ulrich. 2013. 'The Effectiveness of Contracted Coalitions: Private Security Contractors in Iraq'. *Armed Forces & Society*. 39(3):467–488.
- Price, Richard. 1998. 'Reversing the Gun Sights: Transnational Civil Society Targets Land Mines'. *International Organization* 52(3): 613–644.
- Rihoux, Benoît, and Charles C. Ragin, eds. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks: Sage.
- Russett, Bruce, and John R. O'Neal. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organizations*. New York: Norton.
- Salter, Mark B., and Can E. Mutlu, eds. 2013. *Research Methods in Critical Security Studies: An Introduction*. Abingdon: Routledge.
- Schatz, Edward. 2009. *Political Ethnography: What Immersion Contributes to the Study of Power*. Chicago: University of Chicago Press.
- Schmid, Alex P., ed. 2013. *The Routledge Handbook of Terrorism Research*. Abingdon: Routledge.
- Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.
- Shepherd, Laura J., ed. 2013. *Critical Approaches to Security: An Introduction to Theories and Methods*. Abingdon: Routledge.
- Spencer, Alexander. 2010. *The Tabloid Terrorist: The Predicative Construction of New Terrorism in the Media*. Houndmills: Palgrave Macmillan.
- Sprinz, Detlef F., and Yael Wolinsky-Nahmias, eds. 2005. *Models, Numbers, and Cases: Methods for Studying International Relations*. Ann Arbor: University of Michigan Press.
- Stump, Jacob L., and Priya Dixit, eds. 2013. *Critical Terrorism Studies: An Introduction to Research Methods*. Abingdon: Routledge.
- Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca: Cornell University Press.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. New York: McGraw-Hill.
- Wardlaw, Grant. 1982. *Political Terrorism: Theory, Tactics and Counter-Measures*. 2nd edn. Cambridge: Cambridge University Press.
- Williams, Paul D. 2013. *Security Studies: An Introduction*. 2nd edn. London: Routledge.
- Yanow, Dvora, and Peregrine Schwartz-Shea, eds. 2006. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Armonk: M.E. Sharpe.

2 Conceptualizing political violence of non-state actors in international security research

Andreas Armborst

Introduction

Studying non-state actors in international security research can be a daunting challenge. So much so that it takes a whole book, such as this one, to discuss all the potential research strategies. This area of research is difficult for many reasons, and these reasons can be related to methodology (e.g. sampling, collection and analysis of data), theory (e.g. competing paradigms – see Walt 1998: 38), and conception and definition.

This chapter illustrates issues of conceptualization based on the question of how to conceptualize terrorists – who represent a particular type of non-state actors – and, more specifically, how to conceptualize their violent actions. In social research, one has different options for operationalizing terrorism and political violence, each of which has its own methodological, theoretical and conceptual advantages and flaws. The goal of this chapter is to enable the researcher to make a conscious decision about which approach is most suitable and to explain how one can operationalize particular research objectives when studying non-state actors in the realm of international security.

As the most prevalent form of violent conflict today, clashes between non-state actors, as well as between non-state and state actors, have replaced the classic inter-state wars of the 20th century (Kearns and Young 2014: 256). Almost every contemporary armed conflict involves a wide range of participants, including private military and security companies (PMSCs) (see Chapter 4 by Magnus Dau and Marlen Martin, Chapter 6 by Joakim Berndtsson and Chapter 8 by Andreas Kruck), private militias, rebels (see Chapter 3 by Alexander Spencer), defected battalions, pirates, warlords, organized criminal gangs, cartels, insurgents, self-proclaimed states and pro-this or anti-that groups. ‘Classic’ geopolitical disputes, such as the South China Sea, and inter-state wars, such as the conflict between Russia and Georgia in 2008, seem to be the last of their kind.

In order to understand contemporary armed conflicts and international security, it is useful to have a thorough grasp of the nature of violence. Violence is always employed as a means to an end. Different actors use political violence, including terrorism, for different purposes. Because a non-state actor can be

anyone who does not represent a sovereign state, the motives of these individuals are highly diverse. Non-state actors are driven by a wide variety of interests, including commercial, ideological, political, religious, ethnic, racist, sexual, narcissistic, egotistic, megalomaniac, nepotistic or tribal – and often by a combination of them. Consequently, it is not surprising that the theoretical explanations for political violence committed by non-state actors differ from those for violence committed by states, which usually act (or claim to act) in the interest of national collectives, such as to maintain sovereignty and security or to secure access or rights to natural resources and energy.

State actors may establish a monopoly on the legitimate use of force; however, even totalitarian states find it virtually impossible to monopolize *the* use of force. Often, terrorists apparently want to demonstrate just this. For some groups, killing indiscriminately is the preferred choice to demonstrate how powerless the regime actually is.

Because political violence is such a fundamental element of international security, it is worthwhile taking a closer look. Here, I present different conceptual avenues for the study of terrorism and political violence as utilized by non-state actors. The sections that follow may be seen as a beginner's guide to terrorism research. They demonstrate how to utilize five conceptual approaches to the study of terrorism and political violence: the *constructivist approach*, in which terrorism is just a political claim without an empirical manifestation in the real world; the *criminological approach*, which holds that terrorism is a type of deviant behaviour and therefore can be explained through criminological theory; the *legal approach*, which is based on the assumption that terrorism is any action that is legally defined as such; the *academic consensus approach*, which enumerates as many definitional characteristics of terrorism as necessary to reach a certain consensus about its definition regardless of how broad such a definition is; and the *gradual modelling approach*, which applies the concepts of choice, preference and utility from economic theory to differentiate the types of targets selected in political violence.

When one begins to study the political violence committed by non-state actors in international security, it is important to be aware of how complex this topic is. Many newcomers to the field who initially wanted to study terrorism as a type of behaviour and political action end up studying something else. This chapter provides an overview of the different directions terrorism research can take so that the decision to undertake such studies is a deliberate one. Studying terrorism poses a difficult challenge, not because of what the terrorists do, but because of the complex discourses surrounding this topic.

Utilizing different conceptual frameworks of terrorism for research

There is by no means a shortage of concepts and definitions of terrorism, and they each have advantages and weaknesses when they are applied in empirical research. For example, a researcher who adopts a constructivist approach within

international relations (Buzan *et al.* 1998) may investigate how and with what consequences political players use the label ‘terrorists’ to frame certain groups as a threat to domestic, national or international security. The data and methods used in this approach differ from those used in a research project on, say, the ethnography of Salafi insurgent groups in Chechnya. The choice of concept, data and methodology all depend on the researcher’s objective. The following sections describe five conceptually different starting points for the study of terrorism.

The constructivist approach

According to the radical constructivist approach, terrorism is just a ‘political claim’ (Hayes 2015: 3) that has no other empirical manifestation. Hayes’ assertion is based on the fact that terrorism is indeed an ‘essentially contested concept’ (Weinberg *et al.* 2004; Gallie 1956). The problem with the current conceptual model of terrorism is that there is no universally shared model to begin with. Politicians, international organizations, academics and the media all appear to have slightly different – sometimes even contradictory – views when answering the seemingly simple question ‘What is terrorism?’.

In light of this conceptual anarchism it seems logical to adopt the radical constructivist view and say that terrorism is just a political claim with no additional empirical meaning. Constructivist research in international security studies investigates ‘the impact of ideas’ (Walt 1998: 40), such as the ‘idea of terrorism’, or Olof Palme’s revolutionary ‘idea of common security’ (Wiseman 2005). Given that ideas can be powerful drivers of social and political change, it is worthwhile researching them in depth. Constructivist research on terrorism can also offer a historical perspective on the emergence of this – essentially contested – concept and the historical consequences of its introduction into the realm of international security (e.g. the War on Terror). Often, interesting yet barely analysed primary research material is available to address constructivist questions, such as protocols and reports of international commissions (Thakur *et al.* 2005), governmental statements and court papers.

However, the constructivist approach to research on terrorism distracts from the behavioural aspect, namely that terrorism is a particular act of political violence. Given the arbitrary usage of the term ‘terrorism’ in common parlance to describe actions and actors that have nothing in common except that they are associated with political violence of some sort, it is indeed worth asking whether the concept of terrorism has any analytical usefulness at all or whether it may even obstruct empirical research into political violence: as Hayes suggests, ‘the term terrorism creates the false impression that the action it describes represents *a special or unique phenomenon*’ (Hayes 2015: 1, emphasis added). However, later in this chapter, I will answer the question of whether the concept of terrorism is still a useful one by arguing: yes it still is, but not in its current form. In order to conceptualize terrorism as a distinctive type of political violence, we can tackle the problem from the opposite direction, that is, not from the terminological side but

from the empirical side – by asking whether there is a *special or unique phenomenon* within the behavioural range of political violence that we can conceptualize as such, thus leaving aside the forever-contested question of whether ‘terrorism’ is actually the right word.

Conceptualizing terrorism as a crime?

Another intuitive idea is to conceptualize terrorism as a crime. Terrorism apparently is a crime, and criminological theory provides a whole variety of explanations for criminal behaviour. It therefore seems reasonable to capitalize on criminology, which is the discipline of ‘law-making, law-breaking, and reactions to law-breaking’ (Sutherland 1934: 3). Criminology has an affinity with terrorism studies, so it is not surprising to see criminological theory applied to the study of political violence (see the 2015 special issue of *Terrorism and Political Violence* entitled ‘Criminological Theory and Terrorism’, edited by Freilich and LaFree 2015). However, criminology also struggles with the topic of terrorism and political violence, because when one conceptualizes terrorism as a type of criminal behaviour, three anomalies arise. It is worth reflecting on these anomalies because they tell us a great deal about the nature of terrorism.

- 1 *Ordinary criminals break the law; terrorists abolish the law.* Unlike perpetrators of stereotypical crimes such as theft, vandalism, drug offences or tax fraud, terrorists are convinced that rather than breaking the law, they are restoring justice. From the terrorist’s point of view, terrorism is not a crime but the reaction to a crime (Black 2004). This is of course true for all forms of retaliation, retribution, criminal self-help, street justice and vigilantism (Black 1983). People may take the law into their own hands when they think penal measures are not sufficient or not effective in addressing their grievances. Essentially, vigilantism and street justice are substitutes for penal measures. Yet, criminal justice, street justice, vigilantism and terrorism are all different things. Criminal punishment is a lawful response to an unlawful act; vigilantism and street justice are unlawful responses to an unlawful act; but terrorism is an unlawful response to a lawful status quo. The grievance of the terrorist is not a legally protected good – terrorists do not look for substitutes to enforce the law, they are trying to change or abolish the law.
- 2 *Ordinary violent offenders prefer particular victims; terrorist offenders are indifferent in the choice of victims.* Roughly speaking, ordinary criminal violence is either a means of coercion or a means of retaliation. Usually people resort to violence either for moralistic ends, such as to retaliate against an actual or perceived injustice, or as a form of coercion to satisfy monetary, sexual or other desires. In all these cases, perpetrators are never completely indifferent or indiscriminate when choosing their victims. Of course, when targeting a victim, a robber has more options than does someone who wants to take revenge on a thief for stealing his smartphone. Terrorist offenders, however, are almost completely indifferent in the choice

of victims because they seek attention. And to attract attention, it does not matter who is killed, as long as the perpetrators *claim* that their victims are somehow responsible for the grievances that motivated the killing. Whereas the individual victims are random, the location of the attack is usually not. Terrorists prefer soft, symbolic targets so they can inflict a maximum number of casualties with minimal effort and weaponry. Because terrorists kill indiscriminately, they can credibly threaten entire populations, albeit to a limited extent. However, just as the victim is exchangeable, so is the perpetrator. It does not matter who actually carries out the terrorist attack, as long as a group can credibly claim responsibility for it. The constellation between the perpetrator of terrorist violence and the victim(s) is characterized by replaceability on both sides: the sending end and the receiving end of violence. This unlikely offender–victim constellation is best illustrated by the lone-wolf terrorist who kills a person he does not know on behalf of a person or group he does not know either. Classic criminological theory cannot explain such behaviour.

- 3 *In response to ordinary criminal behaviour, the state applies criminal law; in response to terrorism, the state changes criminal law.* The state's reaction to terrorist crime is the third anomaly. States throughout the world have violated their own rule of law when fighting against terrorism. These violations include preventive and pre-trial detention, extraordinary rendition and the torture of terrorist suspects (Sonderegger 2012), to name the most obvious examples. Ordinary crimes, regardless of how despicable or how enormous, do not have as great an impact on the criminal justice system. Terrorism has altered the legal doctrines of states by shifting them towards more preventive approaches (Müller 2011), and some states respond to non-state terrorism with state terrorism.

These three anomalies show that criminology has a conceptual problem when it comes to terrorism. Terrorism is a crime, but it is fundamentally different from other types of deviant behaviour. Yet it is precisely these differences that can tell us a great deal about the peculiarities of terrorism.

Utilizing legal definitions of non-state actors and terrorism

Another way to conduct terrorism research is to use legal definitions. It can be convenient and pragmatic to operationalize terrorism according to one of the many legal definitions that have been proposed in the past (Young 2006). For example, there are 13 United Nations Conventions on Terrorism that define certain terrorist acts, such as the taking of hostages, bombings, terrorist financing and nuclear terrorism. The criminal codes of different nations differ greatly from one another. The United Kingdom's Terrorism Act 2000, for instance, defines terrorism as:

The use or threat of action designed to influence the government or an international governmental organization or to intimidate the public, or a section

of the public; made for the purposes of advancing a political, religious, racial or ideological cause; and it involves or causes: serious violence against a person; serious damage to a property; a threat to a person's life; a serious risk to the health and safety of the public; or serious interference with or disruption to an electronic system.

Legal definitions provide a certain degree of consensus and are legal facts within the domestic and international political context. There is nothing wrong with utilizing legal definitions for research as long as such use aligns with the general research objective. In particular, constructivist studies can make use of legal definitions because such definitions are a concrete example of how powerful political actors construct terrorism. But it is also important to keep the broader perspective in mind: if terrorism research is predominantly based on legal instead of scientific operationalizations, its results will be systematically biased because the research community then studies terrorism as a legal category and not as a particular act of political violence.

In addition, legal and academic definitions of terrorism can differ because they serve different purposes within their respective domains. Scientists operationalize terrorism to identify its causes and consequences. Legislators operationalize terrorism to prosecute and ultimately prevent it. Scientific and legal definitions operate in different professional areas. Criminology textbooks teach that crime is a tremendously contingent and artificial category of human behaviour: 'The society allocates significant financial, human and institutional resources just to determine which concrete behaviour ought to be considered delinquent or criminal' (Sack 1968: 442, author's translation). The same applies to terrorism. Legal definitions of terrorism are artificial constructs, resulting from complex social negotiations. A legal definition of political violence is often of limited value to the social sciences (in the same way as academic definitions are often of limited values for legislators), because both domains have different purposes – that is, criminal prosecution and knowledge production, respectively. Of course, the constructivist view is that scientific knowledge is just as socially negotiated as legal categories are. Be that as it may, judges and prosecutors, like scientists, must collect evidence to decide whether a given act was indeed an act of terrorism, or whether or not a given group is to be designated a terrorist group. But this evidence is evidence according to the social construction of law, and not evidence according to (the social construction of) science. In many cases, both types of evidence correspond (e.g. forensic evidence), because both judges and scientists seek 'the truth'. But it would be flawed to assume that these fact-finding regimes always come to the same conclusion in classifying violence as terrorism. For example, in the United Kingdom, serious violence against property can be an act of terrorism, whereas political sciences usually consider terrorism to be violence against people. No one single interpretation is right or wrong in itself, but both professions would be misguided if they relied exclusively on the other's evidence and definitions. A legislator would fail to designate groups that are indeed terrorists as terrorist groups (according to the

government); a judge would fail to convict a terrorist who is indeed a terrorist (according to the law); and a researcher would operationalize groups that are not actually terrorists as terrorist groups (according to social science).

Utilizing enumerative definitions of terrorism

Another popular approach in academic research is to utilize an enumerative definition of terrorism, as formulated by Schmid (2011), or to adopt certain parts of this definition. Like legal definitions, this definition includes objective and subjective characteristics of violent behaviour that qualify as terrorism. The purpose of an enumerative definition is to reach a certain degree of consensus. It has the advantage that it has undergone rigorous scrutiny by the academic community and that it is less likely to change than legal or political definitions, e.g. after a legal reform. As such it describes terrorism in a more neutral manner as a certain *modus operandi* of political violence rather than as a punishable act or as a derogative political claim.

Because terrorism is an ‘essentially contested concept’ (Weinberg *et al.* 2004; Gallie 1956), it took the academic community 40 years to formulate some sort of consensus definition, and this does not seem to be the end of the story. But why is it so difficult to define a certain type of political violence as terrorism? Equally surprising is why people are still trying to do so. The standard answer to the former question is because ‘one man’s freedom fighter is another man’s terrorist’. No matter how often this slogan is repeated, it remains misleading insofar as it neglects the core of the issue, namely which variations of political violence constitute terrorism and which do not (Richardson 2006: 26). The slogan illustrates only one point: that in political conflicts there will always be different opinions about what constitutes legitimate cause. The question of whether or not terrorists have legitimate cause is necessarily a subjective assessment and therefore not suitable as a definitional criterion. For this reason, the enumerative definition avoids any explicit moral assessment of whether a given cause is legitimate or illegitimate. This is a major analytical advantage; however, it introduces an implicit moral assessment through the backdoor by stating that terrorism deliberately targets (innocent) civilians, which again makes terrorism *illegitimate by definition*. This is not to deny that a deliberate attack against civilians is illegitimate according to almost all ethical standards and common sense. But, at the same time, the term ‘civilian’ admits the concept of innocence and guilt into the definition of terrorism and thereby reopens the door to controversial debate.

Another weakness of the definition is that it requires quite an extensive enumeration of characteristics to reach a sort of consensus. The ‘revised academic consensus definition of terrorism’ (Schmid 2011) includes 12 definitional features. If an action or behaviour is of such particularity that it requires the enumeration of 12 checklist-like features, we must ask whether we are in fact dealing with an empirically observable phenomenon or with an academic artefact. Moreover, such a bulky definition is cumbersome to operationalize in a

given research design. Just imagine the troubles of formulating and testing a theory that explains the following:

Physical violence, or the threat thereof, *carried out* in single-, dual- or multi-phased sequences of action, *intended* to terrorize, intimidate, antagonize, disorient, destabilize, coerce, compel, demoralize or provoke a conflict party *by targeting* civilians, non-combatants or other innocent and defenceless persons, *perpetrated by* small groups, diffuse transnational networks, as well as by state actors or state-sponsored clandestine agents.

The ‘consensus definition’ is – necessarily – so vague as to level all potential points of disagreement.

On reviewing different academic definitions of terrorism, Andrew Silke concluded: ‘The various definitions reach from the absurdly over-specified to the unacceptably over-general’ (Silke 2004: 3). Given that the ‘consensus definition’ by Schmid appears to be of little analytical use owing to its bulkiness and vagueness, it is reasonable to identify those definitional parameters that are crucial for conceptual clarification. A pragmatic approach to this problem would be to utilize only certain elements from the enumerative definition (which of course may come at the expense of consensus). The developers of the Global Terrorism Database (GTD), for example, allow users to filter the 140,000 attacks it has on record with respect to ‘only’ three definitional criteria.¹

Another reason why the definitions of ‘terrorism’ continue to be challenged is that they regard terrorism as a mutually exclusive category within the spectrum of political violence. But such exclusiveness is questionable. Distinguishing terrorist violence from other types of political violence is difficult because the boundaries between them are blurred. The next section describes a gradual model of terrorism and political violence as a fifth option of how to conceptualize political violence of non-state actors. Its purpose is to avoid some of the challenges presented by the legal and enumerative definitions.

Utilizing a gradual model of political violence

In view of the previous discussion of the strengths and weaknesses of conventional definitions, it is clear that research on terrorism is trapped in a conceptual deadlock. Analysis of legal definitions according to a constructivist approach is a viable option for avoiding the definitional challenges, but it does not tell us much about the violence itself. In addition, enumerative definitions of terrorist violence imply that terrorism is an illegitimate form of political violence because it indiscriminately and deliberately targets (innocent) civilians. It is for this reason that terrorism still remains an essentially contested concept.

But how can we exclude implicit moral assessments when defining a behaviour that is *de facto* and *de jure* illegitimate? First, we can take for granted that terrorism is illegitimate. Second, we have to replace the contested definitional element ‘deliberately and indiscriminately targeting (innocent) civilians’.

If we accept the ‘consensus definition’, it becomes clear that the essential characteristics of terrorist violence are somehow determined through the configuration of three variables: target, expectation and goal of the perpetrator. When conceptualizing terrorist violence, the question of whether the target is a civilian or a combatant, innocent or guilty, neutral or partisan, does not matter. What does matter is the degree to which the perpetrator is *indifferent in his or her choice* between civilian and combatant, innocent and guilty, neutral and partisan targets.

A gradual model of terrorism states that political violence is terrorist to the degree to which the perpetrator is indifferent in the choice of targets. This understanding is based on the rather casual observation that ‘It appears as though terrorists rarely have a very coherent idea of what kind of reaction they will get [...]. They can countenance opposite reactions, from capitulation to widespread repression, and be almost equally pleased’ (Richardson 2006: 128, 131).

This observation also answers the question posed at the outset: is there ‘*a special or unique phenomenon*’ within the behavioural range of political violence that we can conceptualize as such? Violent actors that are indifferent in their choice of targets are a special and unique phenomenon indeed.

The gradual model of terrorism utilizes Max Weber’s concept of ‘an ideal type’ (Weber 1968: 191). An ideal type, also known as a pure type, is an exaggerated description of an empirical phenomenon (the ‘real’ type); it is exaggerated to illustrate the defining characteristics of a given social phenomenon. The ideal type functions as a conceptual fixing point to which empirical observations of political violence can relate. Applied to the model of political violence, this means that there are probably no actors who are completely indifferent in the choice of targets and in their expectations about what will be achieved through an attack. The crucial point is that some actors are more indifferent than others and that these differences allow us to distinguish different forms of political violence (Armbrorst 2010). Whether or not ‘terrorism’ is the correct term to describe political violence at the left end of the continuum is another question; for lack of a better term, I shall stick with ‘terrorism’ here.

Figure 2.1 illustrates the gradual model. At the far left end of the continuum (terrorist violence), actors equally prefer all potential targets because they assume they will benefit from whatever the consequences of an indiscriminate attack. At the right end of the continuum (assassination), the preferred target is limited to one, because there are strictly limited expectations as to what will be achieved by attacking this specific target. No other target would be useful to the perpetrator. Other types of political violence oscillate between these two poles of target preference. The utility curve shows how the expected benefits of the attack vary between different targets. In the case of terrorism, the perpetrators are equally satisfied with all potential targets because the perpetrators have no specific expectation about what will be accomplished – their only expectation is that they will benefit from the consequences of the attack in some way. To put it figuratively, terrorists think that they have been dealt a bad hand in political poker. By causing havoc, they are hoping for a new deal and a better hand. If this does

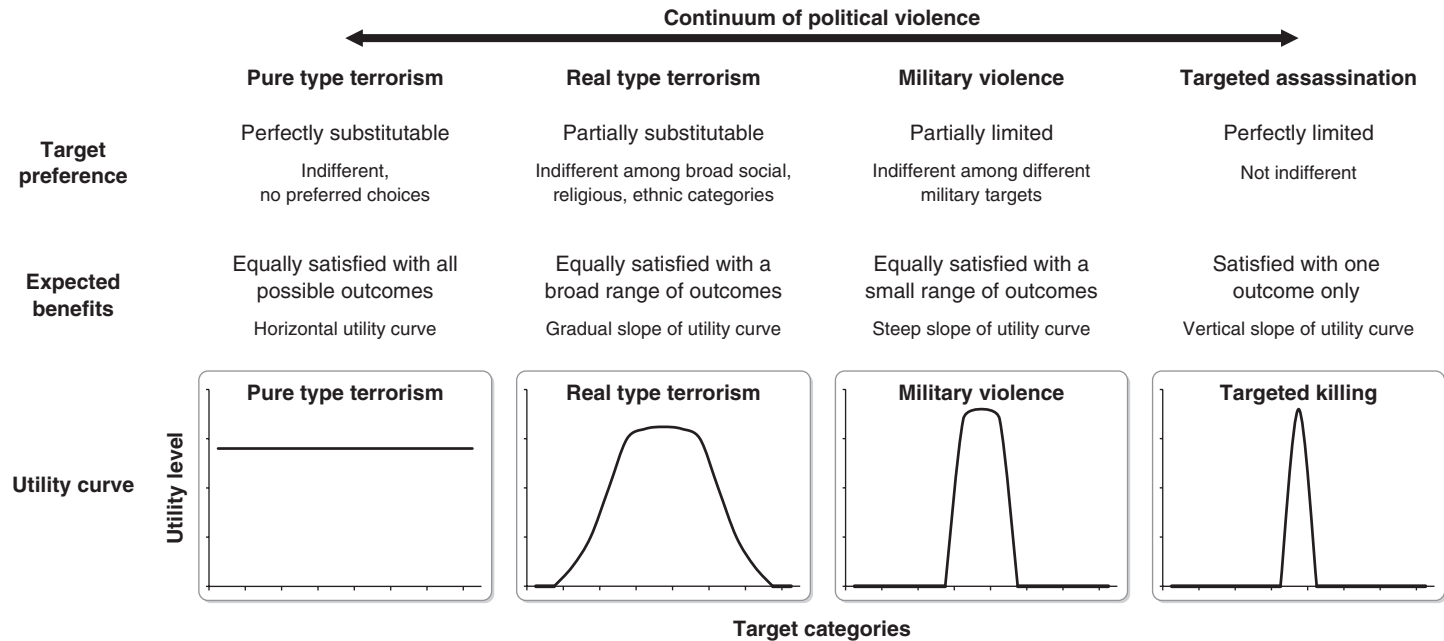


Figure 2.1 Gradual model of political violence.

Source: figure and text revised, based on Armbrist 2010: 425.

not happen, they can strike again. The assassination strategy, by contrast, has a specific target and a specific purpose. Some militant groups employ both tactics or any militant manoeuvre that is between these two poles.

From conceptual models to research scenarios

It is beyond the scope of this chapter to provide detailed research scenarios for each of the five conceptual approaches just described, but I want to offer some general ideas about how to integrate the gradual model of political violence into a research design. This exercise also serves to illustrate how research designs, and ultimately research outcomes, are shaped by the conceptualization of terrorist violence. According to the gradual model, *political violence is designated terrorism depending on the degree to which the perpetrators are indifferent in the choice of their targets*. Perpetrators show no preference when deciding between two or more targets if the targets have the same utility – that is, if the expected benefits of the attack are equally preferred by the perpetrator. The terminology of this definition is borrowed from economic theory, and although it may seem cynical to compare the choices of terrorists with consumers' buying decisions, the crucial point is that militant actors *can* be more or less indifferent in their choice of targets. Potential research questions that originate from this conceptualization include the following: What makes militant actors more or less indifferent in their choice of targets? What sociological, psychological, strategic and political circumstances increase or decrease the perceived utility of different categories of targets for militant actors? Violence itself has no utilitarian value; it always appears in combination with something else, such as strategic or tactical expectations, retaliation, ransom demands or hate. These expectations should be considered along with the *modus operandi*.

In a classic research design, the continuum of political violence can take the place of the dependent variable (the variable to be explained by other variables) with an ordinal scale (rank order). This variable can be included in qualitative, quantitative and analytically mixed research designs. To construct it, some information is required about the actor's subjective expectations about engaging in political violence. Several types of documents and data can provide this information, such as claims of responsibility in which the perpetrators describe their reasons for the attack; ideological statements that describe the rationale behind the use of force; psychological assessments; face-to-face interviews with convicted terrorists; and statistical and documentary databases such as the GTD, the WikiLeaks Iraq War Logs with its nearly 400,000 SIGACT reports (Significant Action of War) and the Afghan War Logs with its 91,000 incident reports.

Whatever the database for the operationalization, it is crucial to develop transparent and reasonable criteria for ranking the cases according to their degree of indifference. For example, if the data include a sample of 250 incidents of political violence, it must be clear why the researcher classifies one incident as an indication of 'high target preference' and another as an indication of 'low target preference'. (The same applies when individuals or groups are ranked.) Another

practical question is what scale or units should be chosen to construct the variable. This choice also depends on the data being used. If the content of the data is nuanced and rich – that is, if it contains sufficient details about the preference level and (expected) utility of the attack – it may justify a scoring system from 1 to 100. In other cases, the data might not include sufficient indications to differentiate political violence to such a degree. Thus, the dependent variable would have fewer categories, such as high, medium and low. In any case, the researcher must make a justifiable and transparent decision on how to operationalize the variable.

In the next step, the researcher can look for correlates of political violence to determine variables that are associated with a high or low degree of target preference. For example, what particular political or psychological circumstances systematically co-occur with high levels of indifference of militant actors? This question can be addressed both by qualitative and by quantitative analysis. In content analysis, for example, the researcher can investigate whether a high (or low) degree of target preference is associated with particular narratives or discourses, and if so, what they are. Statistical analysis can be used to investigate what variables explain differences in target preference and how they interact. The applied statistical model – for example, whether linear regression, generalized linear models or log-linear models must be used – depends on the level of measurement. It is also crucial to consider whether the data being used are on the aggregate level (as with country indicators such as poverty, unemployment and crime rates) or on the individual level (with single measurement points – e.g. the income of every single individual in a given country – instead of the average income). The aggregate level of the data has far-reaching consequences for subsequent statistical analyses, which cannot be addressed here.

From the discussion in this chapter, one can conclude that there are different ways and means of conceptualizing political violence committed by non-state actors. The bad news is that the study of terrorism and political violence is a messy and disparate field of research. The good news is that you as a researcher do not have to know every single approach in detail to be able to carry out a good study on this topic. It suffices to know what the different approaches are, what their strengths and shortcomings are and which of them will provide the most useful starting point for a particular research interest or a particular type of data to which you have access or that you want to collect. Of course, this chapter cannot cover modes of data collection and data analysis in all due detail. It rather serves as a point of departure for further empirical research of political violence.

Note

- 1 These definitional criteria are as follows:

The act must be aimed at attaining a political, economic, religious, or social goal [...]. There must be evidence of an intention to coerce, intimidate, or convey some other message to a larger audience (or audiences) than the immediate victims [...].

The action must be outside the context of legitimate warfare activities. That is, the act must be outside the parameters permitted by international humanitarian law (particularly the prohibition against deliberately targeting civilians or non-combatants).

(GTD n.d.)

References

- Armbrorst, Andreas. 2010. 'Modelling Terrorism and Political Violence'. *International Relations* 24(4): 414–432.
- Black, Donald. 1983. 'Crime as Social Control'. *American Sociological Review* 48(1): 34–45.
- Black, Donald. 2004. 'Terrorism as Social Control', pp. 9–18 in *Terrorism and Counter-Terrorism: Criminological Perspectives*, M. Deflem, ed. Amsterdam: Elsevier.
- Buzan, Barry, Ole Wæver, and Jaap de Wilde. 1998. *Security: A New Framework for Analysis*. Boulder, CO: Lynne Rienner.
- Freilich, Joshua D., and Gary LaFree. 2015. 'Criminology Theory and Terrorism: Introduction to the Special Issue'. *Terrorism and Political Violence* 27(1): 1–8.
- Gallie, Walter B. 1956. 'Essentially Contested Concepts'. *Proceedings of the Aristotelian Society* 56(1): 167–198.
- GTD. n.d. 'Criterion'. Available from: www.start.umd.edu/gtd/more/Criterion.aspx (accessed 1 March 2015).
- Hayes, Jarrod. 2015. *Is the Concept of Terrorism Still Useful?* Zurich: ETH Zurich, International Relations and Security Network, Center for Security Studies.
- Kearns, Erin M., and Joseph K. Young. 2014. 'Military Tactics in Civil War', pp. 257–266 in *Routledge Handbook of Civil Wars*, E. Newman and K. DeRouen, Jr., eds. Abingdon: Routledge.
- Müller, Tim N. 2011. *Präventive Freiheitsentziehungen als Instrument der Terrorismusbekämpfung*. Berlin: Duncker & Humblot.
- Richardson, Louise. 2006. *What Terrorists Want: Understanding the Terrorist Threat*. London: John Murray.
- Sack, Fritz. 1968. *Kriminalsoziologie*. Frankfurt am Main: Akademische Verlagsgesellschaft.
- Schmid, Alex P. 2011. 'The Revised Academic Consensus Definition of Terrorism (Rev. ACDT 2011)', pp. 86–87 in *The Routledge Handbook of Terrorism Research*, A.P. Schmid, ed. Abingdon: Routledge.
- Silke, Andrew. 2004. 'An Introduction to Terrorism Research', pp. 1–29 in *Research on Terrorism: Trends, Achievements and Failures*, A. Silke, ed. London: Frank Cass.
- Silke, Andrew, and Jennifer Schmidt-Petersen. 2015. 'The Golden Age? What the 100 Most Cited Articles in Terrorism Studies Tell Us'. *Terrorism and Political Violence*. doi:10.1080/09546553.2015.1064397.
- Sonderregger, Linus. 2012. *Die Rückkehr der Folter? Anwendung von Zwang bei der Vernehmung im deutschen und US-amerikanischen Recht*. Berlin: Duncker & Humblot.
- Sutherland, Edwin. H. 1934. *Principles of Criminology*. Philadelphia: Lippincott.
- Thakur, Ramesh, Andrew F. Cooper, and John English, eds. 2005. *International Commissions and the Power of Ideas*. Tokyo: United Nations University Press.
- Walt, Stephen M. 1998. 'International Relations: One World, Many Theories'. *Foreign Policy* 110: 29–46.
- Weber, Max. 1968. *Gesammelte Aufsätze zur Wissenschaftslehre*. 3rd edn. Tübingen: Mohr.

- Weinberg, Leonard, Ami Pedahzur, and Sivan Hirsch-Hoefler. 2004. 'The Challenges of Conceptualizing Terrorism'. *Terrorism and Political Violence* 16(4): 777–794.
- Wiseman, Geoffrey. 2005. 'The Palme Commission: New Thinking about Security', pp. 46–75 in *International Commissions and the Power of Ideas*, R. Thakur, A.F. Cooper, and J. English, eds. Tokyo: United Nations University Press.
- Young, Reuven. 2006. 'Defining Terrorism: The Evolution of Terrorism as a Legal Concept in International Law and Its Influence on Definitions in Domestic Legislation'. *Boston College International and Comparative Law Review* 29(1): 23–105.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Part I

Interpreting texts



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

3 Rebels without a cause

Narrative analysis as a method for research on rebel movements

Alexander Spencer

Introduction

This chapter is interested in showing how certain understandings of the (international) political world become dominant while others remain marginalized. The method of narrative analysis helps illustrate the dominance of certain stories in public discourse while at the same time indicating the void left by the stories that remain untold or marginal. Narratives and their analysis are important for the social sciences for two interrelated reasons: first, narratives are important for human cognitive processes; humans think about the world and themselves in a narrative form. Second, narratives are important for identity construction as they are present in any form of history-telling from the present to the past in which relationships between the own community and those of others are established.

As the title of this chapter indicates, narrative analysis does, however, not give insights into questions of causation. Rather it offers answers to ‘how possible’ questions and indicates how certain understandings about non-state actors in international security become dominant while others are marginalized. Hence, narrative analysis is especially useful to study how particular (non-state) actors and courses of action (e.g. military intervention or cooperation with non-state armed groups) in international security are legitimized or de-legitimized in discourse. Dominant (de-)legitimizing narratives and their understandings of actors and contexts shape what is considered possible, logical or natural to do – at the expense of alternative discourses and courses of action. In order to show how narrative analysis can be practically used in research on non-state actors in international security, this chapter will outline the method of narrative analysis and apply this to the rebellion and revolution in Libya in 2011. It examines how to reconstruct by means of narratological tools the stories that are told about the conflict in the British media and the realm of the political elite. It will show how the rebels were romanticized (and thus legitimized) in discourse and how these romantic notions of the rebel marginalize other narratives that do not fit this understanding, such as stories about human rights violations by rebels. Narrative analysis was ideally suited to pursue my original research goals, i.e. to reconstruct dominant as well as marginalized understandings of the Libyan rebels in Britain with a view to understanding which British courses of action were made

(im-)possible by these narrative understandings. However, beyond the specific case of Libyan rebels, it could be applied to study a broad range of (contested) security actors (see Kruck and Spencer 2013; Spencer 2014, 2016). It is important to note that the chapter (and the interpretive method of narrative analysis) does not give insights into the material or ideational *causes* of rebellion but rather focuses on the interpretative level of description by showing the dominance of certain stories told by Western (in this case the British) news media and political elite and the marginalization of others.

In pursuit of this, the chapter will be structured as follows: the first part will illustrate the concept of a narrative and how narratives are important for International Relations (IR). Here, a method of narrative analysis is outlined which focuses on the essential narrative elements of setting, characterization and emplotment. The second part will apply this interpretative method to the narratives on rebellion in Libya found in the newspaper media and in parliamentary debates and speeches of leading politicians in Britain. It will thereby show the dominance of a romanticized story of rebellion. The third part will very briefly indicate the marginalization of alternative narratives, which emphasize human rights violations by rebels.

Narratives and their analysis

Concept and relevance of narratives

Narratives can be found in almost every realm of human life where someone tells us about something. Narratives involve a number of very different text types (written, oral and visual) ranging from literary texts such as novels and poems to films, TV reports, newspaper reports and commentary, school and university textbooks, websites and conversations in our daily life (Fludernik 2009: 1). While some simply understand a narrative as ‘telling someone else that something happened’ (Herrnstein Smith 1981: 228), others stress the issue of time, setting and purpose of this telling: ‘Somebody telling somebody else on some occasion and for some purpose(s) that something happened’ (Phelan 2005: 18). Thus, a narrative can be conceived as the (re-)production of an event or of a sequence of non-randomly connected events (see Genette 1982: 127). Moreover, a narrative centrally involves some sort of disruption or specialty, something that makes it meaningful and interesting for a certain audience and thus worth telling. Finally and most importantly for our analysis, narratives contain some sort of actor, i.e. human or human-like agents such as rebels who are characterized in multiple ways (Herman 2002). Combining these elements Michael Toolan considers narratives as “a perceived sequence of non-randomly connected events, typically involving, as the experiencing agonist, humans or quasi humans, or other sentient beings, from whose experience we humans can “learn”” (Toolan 2001: 8).

But how can narrative analysis, commonly the realm of literary studies, help in the understanding of political and social issues in general and rebel movements in particular? Essentially, there are two interrelated lines of argumentation

that one may pursue here: a cognitive perspective and a cultural perspective (Patterson and Renwick Monroe 1998: 315). The first emphasizes that narratives are a 'fundamental instrument of thought [...] indispensable to human cognition generally' (Turner 1996: 4–5). Narratives are considered to be part of human mental activity (Kreiwirth 2000: 305) in the sense that the human brain 'captures many complex relationships in the form of narrative structures' (Fludernik 2009: 1). Similar to metaphors or analogies (Lakoff and Johnson 1980; Spencer 2012), narratives illustrate a cognitive process of making sense of the world through narration. Therefore, analysis of narratives on, for example, rebel movements and their success or failure among a particular audience gives us an insight into common perceptions and thoughts on these actors. The second perspective considers narratives as culturally embedded phenomena that are part of every society. Myths and stories of the past, including stories about political actors such as rebels, are an essential part of all forms of community-building where the constitution of a common identity is sought. Cultural memory is, above all, made up of different narratives that provide 'us with a fundamental epistemological structure that helps us to make sense of the confusing diversity and multiplicity of events and to produce explanatory patterns for them' (Fludernik 2009: 2).

Taking these two perspectives together, one can argue that individuals as well as communities make sense of themselves and of the social world around them through narratives that constitute their identities as well as their understandings of other actors. In that view, rebels may be (co-)authors and 'subjects' of identity-constructing self-narratives as well as objects of media narratives that constitute a particular understanding of rebel movements. By placing oneself, a community or actors such as rebel movements (not necessarily consciously) in a particular narrative and thereby constituting identity, narratives guide action (Somers 1994: 606–607). Although social narratives are very rarely single authored, the analysis of narratives can nevertheless be considered as an investigation into social action and agency. Thus, the analysis of narratives is of relevance particularly for political science as it contributes to our understanding of common conceptions of 'political reality' and widely held attitudes towards political actors such as rebel movements. The analysis of how and why some narratives succeed and others fail to have an impact ultimately helps to understand the distribution and influence of discursive power (Patterson and Renwick 1998: 315–316).

These thoughts are not fundamentally new for political science or IR and are very much in line with the cultural, discursive or post-positivist turn and a constructivist understanding of politics in which reality is constituted through discourse (Brockmeier and Carbaugh 2001: 9). In contrast to a positivist understanding of social science in which there is an effort to establish an objective, if not universal truth similar to the natural sciences, discourse-oriented constructivists highlight the subjectivity of knowledge and deny the existence of an absolute truth. Unsurprisingly, one comes across the concept of narrative most commonly in constructivist and post-structuralist work on, for example, identity

constructions of self and other, and the drawing of borders (Ringmar 1996; Campbell 1998; Hansen 2006; Browning 2008), the influence of (pop-)cultural narratives (Nexon and Neumann 2006; Lacassagne *et al.* 2011) or the use of Hayden White's insights into historical narratives and explanations (Suganami 2008; Roberts 2006). Yet, despite the widespread use of the concept, very little attention has been paid to the roots or rather home turf of narratives, namely the field of Narratology. Overall in IR, the concept of narrative has frequently been used as a synonym for discourse or frame and has not focused on the specific characteristics of what constitutes a narrative. Narratology offers concrete categories for the empirical analysis of narratives through the established consensus of what elements are necessary for making something a narrative (Spencer 2016: 13–44).

Analysing constitutive elements of a narrative: setting, characterization, emplotment

How could one go about conducting analysis of a vast cultural phenomenon such as narratives? The analysis of narratives in literature studies offers a number of categories which help in the investigation. As visible from the above conceptualizations of the term, a narrative is made up of a number of features such as a *scene* or *setting* of a story, an *agent* that is *characterized* as well as an *event* and *emplotment* of the happening in a causal chain (Toolan 2001; Fludernik 2009: 40). All of these elements are elaborated on in the narrative discourse in order to give them a more specific character and a certain evaluative implication. The dimensions of *setting*, *characterization* and *emplotment* can be empirically analysed and are representative of an overall narrative. While not all of these dimensions are always of the same importance, they may all serve as focal points for more detailed analysis.

With regard to the *setting*, the idea is that similar to a stage play or a film the background or location in front of which the story unfolds is of importance for the narrative as a whole. As Toolan (2001: 41) argues: 'The locations [or settings] where events occur are [...] given distinct characteristics and are thus transformed into specific places'. We all want to know where a story takes place and we consciously or unconsciously look for indicators of the surroundings as they give us a clue of the kind of story we are about to indulge in.

A second important narrative element is *characterization* of the agent (see Fludernik 2009: 6). We are all highly interested in what an agent in a story is like. 'The actors are provided with distinct traits. In this manner they are individualized and transformed into characters' (Toolan 2001: 41). Yet it is not so much a one-off description of an agent such as a rebel that is of most importance for the characterization of it, him or her, but the small continuous predications that build an agent into a character (Fludernik 2009: 46). There are a number of ways in which the characterization of an actor in the story can be influenced. The first and the most simple is giving the agent a name or label rather than simply referring to him, her or it by their occupation or role they play in the story. The

giving of a name or a label personalizes the relationship between the reader and the agent in the story. For example calling a rebel a freedom fighter can bring with it a number of associations including the legitimacy of his/her cause. Second, an agent is characterized by being placed in relation to others. For example this can involve hierarchical relationships such as in the family (mother and child) or in society (protector and protected) or more equal relations such as business partners (Fludernik 2009: 44–45).

A third means of characterization involves the description of the agent's physical attributes, such as clothes or the kind of weaponry rebels carry or outer appearance, including facial expression. As most of these features are considered to be a deliberate choice of the agents and under their control they are thought to provide an insight into the character. A fourth possibility of characterizing an agent is through his/her/its thought process or direct speech. What the character thinks or says greatly influences our perception of what that agent is like and how he/she/it becomes a character. While the narrator of a story is fully responsible for suggesting the thought of characters, the direct speech in newspaper commentary or other media channels is not under the control of the narrator (Herman and Vervaeck 2007: 227). Apart from the name giving, the relationship to others, the description of appearance and direct speech, a final very important implicit aspect of characterization is the way in which the agent acts. Ultimately, the behaviour has an important effect on what we perceive the characters such as rebels to be like (e.g. unorganized, brave, freedom loving, determined, etc.).

The third narrative element, including the event and the *emplotment*, are essential for a narrative: in a narrative something has to happen (see Fludernik 2009: 5). What is more, the happening or rather the event understood as an action has to lead to more action. So the events in a narrative do not stand on their own, they have to be placed in relation to each other (Baker 2010: 353). Here we have to distinguish a temporal and a causal dimension in the ordering of events and action. While the temporal elements of a narrative are important as they emphasize or foreground certain events and limit or silence other happenings, the causal dimension, commonly termed 'causal emplotment', elaborates the causal relationship between the elements of a narrative. Emplotment allows us to weigh and explain events rather than just list them, i.e. to turn a set of propositions into an intelligible sequence about which we can form an opinion. The notion of causal emplotment illustrates how events are connected to each other.

Designing and conducting a narrative analysis

How does one practically go about conducting a narrative analysis? A first step would be to identify the subject of interest (here: the rebels in Libya during the uprising against Gaddafi). Second, one has to decide on the realm of analysis, for example, school books, the news media or speeches by the political elite, as well as the time period one is to analyse. This depends very much on the claims the researcher wishes to make. If one hopes to paint a broad picture of an overall perception, one will have to aim to include a wide range of sources representative

of the realm one wishes to represent. For example, when analysing media content, it is important to include the most representative media outlets in terms of size, popularity and range of ideological stances. While these can never be fully representative, one can aim to be as inclusive as possible.

Third, narrative analysis offers the researcher three fundamental categories – setting, characterization and emplotment – which are then used to structure the analysis of the selected material. In the analysis the text material is then dissected into individual words or phrases that fit into the three categories. Following several rounds of dissection, the researcher will start to gain an insight into the dominance of certain representations within each category. This analysis is greatly aided by the use of qualitative data analysis software programs such as *MAXQDA* or *NVivo* which allow the analyst to mark segments of the text and assign categories for each quote. Alternatively, one is able to conduct this analysis using simple marker pens and an Excel spreadsheet to gather the mass of quotes one will encounter for each category.

Fourth, having filled the categories, the analyst identifies the main and dominant elements found in the texts and engages to retell the story by using the quotes as a collage. Of course, in doing so, one can also count the (relative) frequency of certain categories or words in the analysed texts. However, I should stress that narrative analysis is mainly interested in qualitative insights and in reconstructing stories of political actors, settings and behaviour. There are other tools (e.g. quantitative content analysis) which are better suited if one's primary interest is in quantitative evidence that can be found in texts.

As the reader will notice in the example below, rather than simply paraphrasing the content of the different texts, the form of narrative analysis proposed in this chapter relies on the citation of numerous words (adjectives, verbs and nouns) and phrases in order to create a collage from the newspaper reports. This, although at times obstructing the readability due to excessive footnotes, makes the presentation of the narrative elements transparent and verifiable. However, it is important to note that, by creating a collage of a vast number of quotes this chapter is itself a narrative. This narrative is inherently subjective. Yet, through transparency, verifiability and showing the extensiveness of the narrative elements, the subjective story can be made intersubjective.

Narratives of rebellion in Libya

The following section will apply the method of narrative analysis outlined above to the discourse on the rebellion in Libya against Colonel Gaddafi in 2011 found both in the British newspaper media and in British parliamentary debates and speeches of the political elite at the time. I first reflect on the analysis of media narratives before turning to the analysis of narratives of the political elite. Finally, I highlight how narrative analysis understood as a critical method may serve to highlight alternative and marginalized understandings.

Analysis of media narratives

In order to gain an insight into the narratives being told about rebels during the conflict I used the *Nexis* (formerly *Lexis Nexis*) news database to search for the concepts of 'Libya' and 'rebel' in the *Guardian*, *The Times* and the *Sun* newspapers from the beginning of the conflict (17 February 2011) until its official conclusion (31 October 2011). Out of these, the analysis focused on 90 of the – according to the *Nexis database* – most relevant articles. The newspapers were selected as they cover both the left and the right of the political spectrum, as well as a more popular tabloid story about rebellion in Libya. Analysis beyond the 90 articles did not yield any new or alternative narrative elements. In my analysis of the 90 articles, I inductively drew out dominant narrative elements with regard to the setting of the rebellion, the characterization of rebels and the emplotment of the rebellion. In particular, my narrative analysis revealed that romantic elements were (more or less) widespread in the setting, characterization and emplotment of narratives about Libyan rebels.

Focusing on the *setting* depicted in the media narratives, it becomes clear that the narrative of rebellion is set in the wider context of the Arab Spring. However, there are only a few material or scenic indications of a romantic setting that constitute the narrative as one set in an exotic, far away and foreign desert, which is very different to the setting of the everyday in the West. Occasionally one does encounter references and scenes that play in a desert, with 'sand dunes',¹ 'blue sky',² 'palm trees'³ and an 'oasis'⁴ where the 'sun sets'⁵ on rebel positions. However, the exciting and emotional side to the romantic setting appears far more prominently. While the emotion at first includes fear of Gaddafi, this transcends into joy as '[p]eople are not afraid'⁶ anymore. While the civilian population was at first 'deeply fearful of Gaddafi',⁷ the 'fear barrier is broken':⁸

As word spread, Tripoli's streets filled with honking vehicles and soon thousands upon thousands of people were converging on a compound that had previously filled them with dread. Their fear gone, they poured through the broken walls even as bullets and shells whistled overhead.⁹

This emotion of fear is replaced after victories with an 'explosion of joy',¹⁰ 'joyous scenes'¹¹ and 'ecstatic celebrations'¹² as the rebels are enthusiastically embraced and 'greeted by civilians lining streets, cheering and waving rebel flags'.¹³ The rebels are 'cheered',¹⁴ civilians 'flashed the V for victory sign'¹⁵ and residents 'celebrate wildly'.^{16,17} As the narrative element of characterization will show in more detail, the rebel is said to be on the side of the people of the street against an oppressive regime.

With regard to *characterization*, the rebel side of the conflict is commonly labelled as 'revolutionaries'¹⁸ or 'revolutionary forces'.¹⁹ Furthermore, one encounters a number of labels such as 'liberation force',²⁰ 'pro-democracy fighters',²¹ 'freedom crusader',²² 'freedom campaigners'²³ or most commonly 'freedom fighters',²⁴ which link aspects of emplotment with characterization by

including causal motivations and goals underlying the struggle. The individual rebels are characterized not as professional fighters but as ‘young’²⁵ ‘teenager[s]’²⁶ or ‘25-year-old’²⁷ ‘civilians’²⁸ who have taken up arms against Gaddafi. ‘The rebel army is mostly ordinary young people, often well educated, who have had enough of their lives being destroyed by a dictator’.²⁹ Emphasizing their civilian character, many of the depictions of the rebels includes references to their former occupation such as ‘engineer’,³⁰ ‘pharmacist’,³¹ ‘dentist and lab technician’,³² ‘cashier’,³³ ‘house decorator’,³⁴ ‘teacher’,³⁵ ‘student’³⁶ and even an ‘owner of a Liverpool takeaway’.³⁷ Similar to many of the audience reading the story, the rebel is a normal human being one is able to identify with through similarity. The ‘teacher-turned-fighter’³⁸ rebel has ‘no qualifications to be a soldier’³⁹ and is by no means an elite fighter but rather characterized as an underdog in the conflict. The rebels are said to be ‘amateur fighters’⁴⁰ who have a ‘lack of military experience’.⁴¹ Furthermore, the rebels are said to be ‘ill armed’⁴² and ‘ill equipped’⁴³ as they ‘lack the firepower and military discipline’.⁴⁴ Despite this characterization, the rebels remain impressive: ‘There has been a consistent tendency to underestimate the rebel leadership, a tendency to dismiss them as a rabble or a bunch of Islamists. At almost every turn thus far, they have surpassed expectations’.⁴⁵ So they are also considered ‘brave men’⁴⁶ in a ‘heroic uprising against Libya’s murderous tyrant Colonel Gaddafi’.⁴⁷ ‘Most of the revolutionaries are not trained soldiers but the spirit and bravery to fight is there’.⁴⁸ They are fighting for admirable and honourable causes such as their people and homes⁴⁹ or in ‘pursuit of freedom’.⁵⁰ The revolution involves ‘truly inspiring individuals: doctors, teachers, lawyers and engineers willing to die for their freedom’.⁵¹

With regard to *emplotment*, the event of war in Libya is constituted as an asymmetrical conflict in which the righteous weaker rebels are struggling against a stronger and unjust enemy in the form of the Gaddafi Regime. As already indicated in their characterization, the emplotment of the conflict in Libya emphasizes that the ‘rebels were heavily outnumbered and outgunned’:⁵² ‘As the sun set, the rebels were perched behind the few anti-aircraft guns they hold, their thin turrets pointing towards an empty black sky’.⁵³ The conflict is emplotted as a struggle against the odds in which the rebels only have ‘sparse forces’⁵⁴ and ‘Colonel Gaddafi’s forces have better weapons’.⁵⁵ Nevertheless, the underdog rebel forces remarkably win against all odds as ‘we witnessed this volunteer force – armed with a hotch-potch of AK-47s and rocket-propelled grenades – stand firm against a much better equipped army’.⁵⁶

Apart from the asymmetry, the conflict is emplotted as a resistance to an unjust order that ‘terrorised its inhabitants’,⁵⁷ ‘systematically maltreated the civilian population’⁵⁸ and where ‘people suffered from the repression and brutality of Colonel Muammar Gaddafi’.⁵⁹ The unjust order, run by an ‘crazed tyrant’⁶⁰ is constituted as a ‘murderous’,⁶¹ ‘cruel and destructive regime’⁶² responsible for ‘indiscriminate shelling’⁶³ and ‘atrocities’⁶⁴ and ‘the most brutal response to the Arab Spring’.⁶⁵ Therefore, the narrative involves a struggle for an ideal, a fight for liberation and an illustration of love for a country and its people, as rebels are quoted: ‘We love our country. We love each other’.⁶⁶ The narrative constituted

the cause of the conflict to be a 'burning desire for freedom'⁶⁷ and democracy.⁶⁸ The conflict is a 'fight for democracy'⁶⁹ and the rebels 'are fighting for freedom'.⁷⁰ The reader in the UK can identify with the rebels as they 'want freedom and democracy, just like we have in Britain'.⁷¹ 'They are students, engineers, teachers and doctors who came together through Twitter and Facebook. They want freedoms people in Britain and the West take for granted'.⁷²

Overall the British media narrative of rebellion in Libya that I reconstructed from the newspaper articles is a romantic story in an emotional setting in which the rebel is characterized as a young and brave underdog fighting against a brutal and oppressive regime for an ideal such as democracy and freedom.

Analysis of narratives of the political elite

While media discourses are now widely accepted in the social sciences as relevant sites for analysis where political struggles of meaning and power take place, one may still doubt whether these media narratives are also reflected in the political elite narratives of decision makers. Therefore, as a next step after the media analysis, I conducted a narrative analysis of British decision-makers' discourse on the rebellion in Libya. I analysed all the 23 parliamentary debates in the House of Commons on the situation in Libya and on UN Security Council Resolution 1971 on the intervention in the country, between the beginning of the conflict (17 February 2011) until its official conclusion (31 October 2011). This was further supplemented by five public speeches given by Prime Minister David Cameron at the time. Interestingly, my narrative analysis of British parliamentary debates and political speeches revealed that many of the elements with regard to the setting, characterization and the emplotment of the rebel that were found in the media were also clearly visible in the discourse of the political elite. The analysis of political elite narratives thus underlined that certain (romantic) understandings of the Libyan rebel(lion) have become dominant in both British society and policy-making circles.

With regard to the *setting*, one comes across some references that indicate an exotic location that is different to the everyday Briton, by stressing the conflict setting of the Libyan 'desert'⁷³: 'At the meeting Nicholas Sarkozy hosted in Paris, we made the right choice: to draw a line in the desert sand, and to halt his murderous advance by force'.⁷⁴ The conflict is placed in the overall context of the Arab Spring and there are references to the exciting and emotional side of the story, such as celebration, fear, hope and drama,⁷⁵ for example 'This has been the most dramatic episode of what has been called the Arab Spring'⁷⁶ and 'Where there was fear, now there is hope and an optimism and belief that is truly inspiring'.⁷⁷

Many of the narrative elements of *characterization* found in the media discourse are also present in the discourse of the political elite. While one does encounter a number of fairly unromantic labels for the actors involved, such as 'opposition groups',⁷⁸ other labels such as 'the Libyan people'⁷⁹ or 'anti-Gaddafi forces'⁸⁰ indicate sympathy towards them from a British perspective. At the

same time, the narrative includes clearly more romantic labels such as ‘pro-freedom rebels’⁸¹ or ‘free Libyan fighters’.⁸² Yet these fighters are not considered to be professional soldiers but ‘ordinary Libyans from all walks of life’⁸³ who are no match for Gaddafi’s troops:

Ordinary Libyans from all walks of life came together and rose up against Gaddafi. From the villages of the Nafusa mountains to the tower blocks of Misrata, the alleyways of Zawiyah and the streets of Benghazi, the Libyan people fought with incredible courage.⁸⁴

In this struggle, the rebel is characterized as being ‘brave’⁸⁵ and having ‘courage’⁸⁶ as they are ‘fighting the dictator with their bare hands’.⁸⁷ They have successfully ‘liberated themselves’⁸⁸ and are ‘taking their destiny into their own hands’.⁸⁹ Therefore, ‘we should also pay a tribute to the bravery and resilience of the Libyan people themselves. This has been their revolution and none of it could have happened without them’.⁹⁰ Similarly, one is able to identify with the rebels as they are said to be fighting for an understandable cause, such as freedom and democracy. They ‘want the sort of freedoms that we take for granted in this country’.⁹¹

Many of the elements indicated in the characterization of the rebels also reappear in the *emplotment* of the event as a ‘popular uprising’⁹² or a ‘revolution’.⁹³ Again, the conflict is constituted as an asymmetrical struggle against an unjust order where the regime is ‘attacking peaceful protestors’⁹⁴ ‘using the full might of his armed forces, backed up by mercenaries’.⁹⁵ Gaddafi is emplotted as a ‘despot’⁹⁶ and ‘tyrant’⁹⁷ who is ‘cruelly’⁹⁸ committing ‘barbaric acts against the Libyan people’,⁹⁹ ‘murdering’¹⁰⁰ ‘his people indiscriminately’¹⁰¹ with ‘no mercy and no pity shown’.¹⁰² Part of the narrative is emplotted as a struggle for a utopian future in which everything will be much better. The people of Libya are fighting for a ‘better future’,¹⁰³ ‘democratic future’,¹⁰⁴ a ‘peaceful and prosperous future’¹⁰⁵ or a ‘peaceful, successful and democratic future’.¹⁰⁶ Importantly, as already indicated as part of the characterization, the rebels’ struggle against an unjust order is pursued for noble ideals as Libyans are said to be ‘struggling for democracy and freedom’.¹⁰⁷ The desire for ‘openness’,¹⁰⁸ ‘freedom’¹⁰⁹ and ‘democracy’¹¹⁰ are said to be the understandable cause of the conflict with the Gaddafi regime. These values are explicitly linked to British values and are, similar to elements in the characterization of the rebels, a means of understanding and identifying with the righteousness of the struggle for ‘freedoms that people in Britain take for granted’.¹¹¹

Very similar to the narrative elements found in the British media, the story of rebellion in Libya is found to involve an emotional setting in which the young underdog rebel is bravely fighting a far stronger brutal and oppressive Gaddafi regime in an asymmetric conflict for a better, almost utopian, future and the noble ideals of freedom and democracy.

Narrative analysis as a critical method: retracing alternative and marginalized stories (of rebellion)

The critical potential of narrative analysis is its ability to highlight the dominance of certain narratives but also to unmask the silenced discourses of the marginal. Therefore, my narrative analysis also focused on retracing such marginalized narratives. While stories of the rebel who is struggling for noble ideals against an unjust and more powerful order are very much present in both the media and political coverage of the conflict in Libya, other, less amicable narratives only exist at the margins despite their potential truthfulness. The following section will briefly illustrate such a narrative, which tells a story in which rebels are responsible for war crimes and human rights violations during and after the conflict. Of course, it will depend on the goals of the researcher (and the available space in an article, chapter or book) to what extent his or her narrative analysis will put an emphasis on marginalized (rather than or besides dominant) narratives. But from a critical perspective on international security, it is important to focus not only on hegemonic but also alternative stories in one's analysis.

In my analysis of narratives about the rebellion in Libya I found that it is widely accepted that Gaddafi's regime is responsible for a large number of crimes, attacks on civilians and human rights violations; by contrast, stories accusing the rebels of perpetrating such crimes are rare. This does not mean that the rebel side of the conflict simply did not commit such acts, as there are a number of actors who try to emphasize that members of the rebels were in indeed responsible for human rights violations.

For example, Amnesty International stated in its report on the conflict in Libya that, although not on the same scale as the Gaddafi regime, members of the opposition 'have also committed human rights abuses, in some cases amounting to war crimes'.¹¹² Overall, nine pages of the report on the conflict are dedicated to crimes and human rights violations by rebels. Rebels and their supporters are said to 'have abducted, arbitrarily detained, tortured and killed former members of the security forces, suspected al-Gaddafi loyalists, captured soldiers and foreign nationals wrongly suspected of being mercenaries fighting on behalf of al-Gaddafi forces'.¹¹³ Similarly, Human Rights Watch tells a story in which '[r]ebel forces also committed human rights and humanitarian law violations during the armed conflict'.¹¹⁴ Rebels are here accused of 'revenge attacks'¹¹⁵ against suspected Gaddafi loyalists as well as 'arbitrary arrests',¹¹⁶ 'serious abuses',¹¹⁷ 'mistreatment'¹¹⁸ and 'torture in detention, in some cases leading to death'.¹¹⁹ The most prominent case of human rights violations includes accusations of 'revenge killings'¹²⁰ and, in particular, the suspected revenge execution of 53 Gaddafi supporters: 'Some of the victims had their hands bound behind their backs; they all seemed to have been shot at that location'.¹²¹

Few of these stories are taken up in the media¹²² or the political debates in parliament and, at times, the existence of violations by rebels is denied: 'All I can say about the opposition forces and the danger of civilian deaths from their

activities is that, so far, we have no record of their being engaged in attacks on civilians.¹²³ Interestingly, when such stories of situations in which crimes by rebels could be committed, they are emplotted as understandable in the given setting:

Foreign mercenaries, particularly black Africans, have been blamed by rebels for committing the most heinous crimes, including gang rape. It seems astonishing then, in the heat of battle, with a friend just killed, that the rebels did not murder the youth from Chad. Instead, they took him to a hospital and he was treated alongside their friends.¹²⁴

Similarly, with regard to the looting committed by rebels, their acts become understandable as the loot is not really stolen; 'in truth it did not seem much like looting as most of the treasures they carted away were acquired with wealth stolen from the Libyan people.'¹²⁵

In the case of the Gaddafi's killing by rebels, there are a number of voices in the media that emphasize the brutal and criminal nature of the act,¹²⁶ its implications for the human rights situation in Libya¹²⁷ and stress that Gaddafi should have been put on trial:¹²⁸ 'His executioners are as bad as he was, demonstrating complete ignorance of human values'.¹²⁹ Similarly, a number of other actors, such as Amnesty International¹³⁰ and the UN High Commissioner for Human Rights,¹³¹ criticized the death of Gaddafi and called for an investigation.¹³² Hugo Chávez claimed: 'They assassinated him. It is another outrage. We shall remember Gaddafi our whole lives as a great fighter, a revolutionary and a martyr'.¹³³ In contrast, however, the majority of media narratives do not tell the story of Gaddafi's death as a crime or a human rights violation. The killing of Gaddafi is said to be justified as '[t]he murderous tyrant got the fate he deserved'.¹³⁴ 'His death was grotesque and violent. Much as he had lived his life'.¹³⁵ Alternatively, Gaddafi is said to have got off too lightly as '[k]illing him on the spot was an easy way out for him' and that victims 'felt cheated that the dictator had escaped justice'.¹³⁶ Other media reports stress that the killing of Gaddafi was a necessary step to 'help stabilise the position of Libya's new leaders, particularly if his followers now gave up the fight'.¹³⁷ 'Viewed from a "realpolitik" perspective, this was the best possible outcome for the country, its present rulers and those optimistic enough to believe that Libya can eventually be transformed into a vibrant democracy'.¹³⁸ In the story, the person responsible for the death of Gaddafi is not only a killer but a 'Libyan hero'.¹³⁹ 'Why all the fuss about who killed Gaddafi? The man who shot him should be given a medal'.¹⁴⁰ At the same time some of the more critical voices on the killing of Gaddafi are attacked and delegitimized:

Gaddafi has finally got what was coming to him and the first thing out of the UN is, 'We must have an inquiry into exactly how he died'. What did they do to help the good people of Libya overcome this tyrant during the past 40 years? Nothing!¹⁴¹

In contrast to the stories told in the media, some amongst the political elite in Britain, such as defence secretary Philip Hammond, are a 'little' more sceptical about Gaddafi's death: The Libyan revolutionaries' image had been 'a little bit stained' by Gaddafi's death, Hammond told the BBC. 'It's certainly not the way we do things. We would have liked to see Colonel Gaddafi going on trial to answer for his misdeeds.'¹⁴² Similarly, William Hague noted that he would have preferred for Gaddafi to be captured alive: 'We would have preferred him to be able to face justice at the International Criminal Court or in a Libyan court for his crimes. We don't approve of extrajudicial killings'.¹⁴³ However, he goes on to stress that

[a]t the same time we are not going to mourn him. This and the fall of Sirte and Bani Walid is a major opportunity for Libya to be able to move on to what they've fought for all this year, into a free and democratic future.¹⁴⁴

David Cameron was less hesitant as he 'hailed Gaddafi's death as a step towards a 'strong and democratic future' for the north African country'.¹⁴⁵

Despite the existence of crimes and human rights violation by rebels, these stories remained largely on the margins of the narratives being told about rebellion in Libya. Narrative analysis is able to show how some of these alternative narratives on human rights violations by rebels were silenced, ridiculed and explicitly refuted and how other narratives even tried to justify and present these violations as understandable. The analysis of narratives is thus able to uncover dominance and marginalization of political understandings.

Conclusion

This chapter has given an insight into the use of narrative analysis for research into non-state actors in international security. Narrative analysis does not give researchers insights into the causes of rebellion but into the portrayal of how rebellion is told as a story. As the analysis of parts of the British news as well as British parliamentary debates and speeches by the politicians showed, the narratives of rebellion in the Libyan conflict in 2011 were predominantly told in a romantic fashion. This contributed to legitimizing Libyan rebels and British political and military support of these rebels. Many of the narrative elements outlined in pre-existing cultural stories, such as Lawrence of Arabia, including an emotional exotic setting of the desert, heroic and brave young rebels and an asymmetric and almost utopian struggle for an ideal such as freedom, were also present in the narratives on rebellion in Libya. Beyond the mere existence of such romantic story elements the chapter argued that these dominant understandings marginalized and silenced other more negative narratives on crimes and human rights violations despite their potential 'truthfulness'. Narrative analysis can contribute to the understanding and criticism of dominant political perceptions in world politics. It can be applied to study a broad range of contested state and non-state actors and courses of action in international security (Spencer 2016).

Notes

- 1 'Front: A few steps forward, many back on the long road west to Tripoli', *Guardian*, 22 March 2011, p. 1.
- 2 'Just 30 miles from Tripoli, the defiant town that dares to humiliate Gaddafi', *The Times*, 7 March 2011, pp. 6–7.
- 3 'Why 1941 Desert Rats are the key to defeating Mad Dog today', *Sun*, 2 April 2011, pp. 42–43.
- 4 'Libya can stay united after Gaddafi has gone', *The Times*, 31 March 2011, p. 18.
- 5 'Libya: The rebels gather their sparse forces for the battles ahead', *Guardian*, 1 March 2011, p. 14.
- 6 'Libya: "The fear barrier is broken": volunteers flock to join the rebels' drive-in war', *Guardian*, 9 March 2011, p. 16.
- 7 'Front: The graveyards are filling up in Misrata's unexpected war', *Guardian*, 22 April 2011, p. 1.
- 8 'Libya: "The fear barrier is broken": volunteers flock to join the rebels' drive-in war', *Guardian*, 9 March 2011, p. 16.
- 9 'The fall of Tripoli', *The Times*, 24 August 2011, pp. 1, 3.
- 10 'Gaddafi, you are History', *Sun*, 24 August 2011, pp. 8–9.
- 11 Ibid.
- 12 'Don't you know right from wrong?', *Sun*, 22 October 2011, pp. 6–7.
- 13 'He's Buggied Off', *Sun*, 26 August 2011, pp. 6–7.
- 14 'Gaddafi, you are History', *Sun*, 24 August 2011, pp. 8–9.
- 15 'Libya rebels take key city to "start march" on Tripoli', *Guardian*, 15 August 2011, p. 15.
- 16 'Front: Endgame: rebels take Tripoli', *Guardian*, 22 August 2011, p. 1.
- 17 'Libya: Defiant Gaddafi promises he will "fight to the last drop of blood"', *Guardian*, 22 August 2011, p. 4.
- 18 '"We beg the West for a no-fly zone before Gaddafi's jets destroy us all"', *Sun*, 15 March 2011, p. 14.
- 19 'Libya: Rebel push reveals devastating impact of allied air strikes', *Guardian*, 28 March 2011, p. 4.
- 20 '4p Off Petrol if Gaddafi Goes', *Sun*, 23 August 2011, pp. 6–7.
- 21 '30 killed by Gaddafi's mercenaries in bloody battle for control of rebel-held stronghold', *The Times*, 5 March 2011, pp. 8–9.
- 22 '"We beg the West for a no-fly zone before Gaddafi's jets destroy us all"', *Sun*, 15 March 2011, p. 14.
- 23 'Blast Days of Gaddafi', *Sun*, 29 March 2011, pp. 12–13.
- 24 'Rebels struggling to flip the power switch despite a detailed blueprint', *The Times*, 25 August 2011, pp. 6–7.
- 25 'Front: Libya: "They are professional soldiers. And they are guarding someone"', *Guardian*, 26 August 2011, p. 4.
- 26 'Gaddafi, you are History', *Sun*, 24 August 2011, pp. 8–9.
- 27 'Rebels close in on Gaddafi', *The Times*, 15 August 2011, pp. 1, 8, 9.
- 28 'Front: The graveyards are filling up in Misrata's unexpected war', *Guardian*, 22 April 2011, p. 1.
- 29 '"We beg the West for a no-fly zone before Gaddafi's jets destroy us all"', *Sun*, 15 March 2011, p. 14.
- 30 'Expats return from UK for rebel assault on Tripoli', *The Times*, 2 July 2011, p. 45.
- 31 'Front: The graveyards are filling up in Misrata's unexpected war', *Guardian*, 22 April 2011, p. 1.
- 32 'Front: Shattered heart of Gaddafi's government', *Guardian*, 25 August 2011, p. 1.
- 33 'Front: Libya: "It is a symbolic victory. Gaddafi is still free"', *Guardian*, 24 August 2011, p. 2.

- 34 'The battle of Qasr al-Haj', *The Times*, 11 June 2011, p. 39.
- 35 'Rebels close in on Gaddafi', *The Times*, 15 August 2011, pp. 1, 8, 9.
- 36 'Apache strikes aim to break deadlock and allow rebel advance on Tripoli', *The Times*, 6 June 2011, p. 6.
- 37 'Expats return from UK for rebel assault on Tripoli', *The Times*, 2 July 2011, p. 45.
- 38 'Rebels close in on Gaddafi', *The Times*, 15 August 2011, pp. 1, 8, 9.
- 39 'It's a buzz fighting Gaddafi ... but I do miss Emmerdale', *Sun*, 10 March 2011, pp. 10–11.
- 40 'Rebels close in on Gaddafi', *The Times*, 15 August 2011, pp. 1, 8, 9.
- 41 'Front: Blow to Gaddafi as Libyan foreign minister defects and flies to UK', *Guardian*, 31 March 2011, p. 1.
- 42 'Muammar Gaddafi: Brutal and unpredictable Libyan leader', *The Times*, 21 October 2011, pp. 74–76.
- 43 'Road to Benghazi beckons for Gaddafi as Libyan rebels retreat under rain of rockets', *Guardian*, 14 March 2011, p. 17.
- 44 'Britain and France "using Qatar to arm Libyan rebels"', *The Times*, 25 June 2011, pp. 16–17.
- 45 'National: Special report: Cameron's war: why PM felt Gaddafi had to be stopped', *Guardian*, 3 October 2011, p. 13.
- 46 'It's a buzz fighting Gaddafi ... but I do miss Emmerdale', *Sun*, 10 March 2011, pp. 10–11.
- 47 '"We beg the West for a no-fly zone before Gaddafi's jets destroy us all"', *Sun*, 15 March 2011, p. 14.
- 48 Ibid.
- 49 'Front: The graveyards are filling up in Misrata's unexpected war', *Guardian*, 22 April 2011, p. 1.
- 50 'The battle of Qasr al-Haj', *The Times*, 11 June 2011, p. 40.
- 51 'Beware. Libya could easily tip over the edge', *The Times*, 30 August 2011, p. 20.
- 52 'The battle of Qasr al-Haj', *The Times*, 11 June 2011, p. 39.
- 53 'Libya: The rebels gather their sparse forces for the battles ahead', *Guardian*, 1 March 2011, p. 14.
- 54 Ibid.
- 55 'Women and children killed in the bloody battle for Zawiya', *The Times*, 10 March 2011, pp. 6–7.
- 56 'I stood rooted to the spot as jet bombed us', *Sun*, 11 March 2011, pp. 10–11.
- 57 'Hit by a Tornado', *Sun*, 28 March 2011, pp. 12–13.
- 58 'Libya: Rebel push reveals devastating impact of allied air strikes', *Guardian*, 28 March 2011, p. 4.
- 59 'The fall of Tripoli', *The Times*, 24 August 2011, pp. 1, 3.
- 60 'Blast Days of Gaddafi', *Sun*, 29 March 2011, pp. 12–13.
- 61 'Where is Gaddafi', *Sun*, 21 March 2011, pp. 4–5.
- 62 'Muammar Gaddafi: Brutal and unpredictable Libyan leader', *The Times*, 21 October 2011, pp. 74–76.
- 63 'Front: The graveyards are filling up in Misrata's unexpected war', *Guardian*, 22 April 2011, p. 1.
- 64 'Don't you know right from wrong?', *Sun*, 22 October 2011, pp. 6–7.
- 65 'National: Special report: Cameron's war: why PM felt Gaddafi had to be stopped', *Guardian*, 3 October 2011, p. 13.
- 66 'Libya: "The fear barrier is broken": volunteers flock to join the rebels' drive-in war', *Guardian*, 9 March 2011, p. 16.
- 67 'Mad Max v Mad Dog', *Sun*, 9 March 2011, pp. 10–11.
- 68 'Britain's message of solidarity with the fight for democracy in Libya', *The Times*, 22 August 2011, pp. 6–7.
- 69 Ibid.

- 70 'Libyators', *Sun*, 28 February 2011, p. 19.
- 71 'It's a buzz fighting Gaddafi ... but I do miss Emmerdale', *Sun*, 10 March 2011, pp. 10–11.
- 72 Ibid.
- 73 William Hague, Statement in the House of Commons, 7 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110307/debtext/110307-0001.htm#11030711000002 (accessed 4 March 2014).
- 74 David Cameron, Speech at the London Conference on Libya, London, 29 March 2011. Available from: www.newstatesman.com/2011/03/libya-help-humanitarian (accessed 4 March 2014).
- 75 Luciana Berger, Statement in the House of Commons, 28 February 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110228/debtext/110228-0001.htm#11022819000002 (accessed 5 March 2014).
- 76 David Cameron, Speech to the UN General Assembly, New York, 22 September 2011. Available from: www.newstatesman.com/global-issues/2011/09/arab-world-region-libya-act (accessed 4 March 2014).
- 77 Ibid.
- 78 Hague, Statement in the House of Commons, 7 March 2011.
- 79 David Cameron, Statement in the House of Commons, 5 September 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110905/debtext/110905-0001.htm#1109054000001 (accessed 4 March 2014).
- 80 Douglas Alexander, Statement in the House of Commons, 15 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110315/debtext/110315-0001.htm#11031569000027 (accessed 4 March 2014).
- 81 Peter Bone, Statement in the House of Commons, 19 June 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110719/debtext/110719-0001.htm (accessed 4 March 2014).
- 82 David Cameron, Statement on Libya, 10 Downing Street, 22 August 2011. Available from: www.newstatesman.com/politics/2011/08/arab-libyan-step-today-support (accessed 4 March 2014).
- 83 David Cameron, Statement in the House of Commons, 21 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110321/debtext/110321-0001.htm#1103219000002 (accessed 5 March 2014).
- 84 David Cameron, Statement in the House of Commons, 5 September 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110905/debtext/110905-0001.htm#1109054000001 (accessed 4 March 2014).
- 85 Cameron, Speech to the UN General Assembly, New York, 22 September 2011.
- 86 Nicholas Soames, Statement in the House of Commons, 5 September 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110905/debtext/110905-0001.htm#1109054000001 (accessed 4 March 2014).
- 87 Ann Clwyd, Statement in the House of Commons, 28 February 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110228/debtext/110228-0001.htm#11022819000002 (accessed 5 March 2014).
- 88 Cameron, Speech to the UN General Assembly, New York, 22 September 2011.
- 89 David Cameron, Statement in the House of Commons, 18 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110318/debtext/110318-0001.htm#11031850000007 (accessed 5 March 2014).
- 90 Cameron, Statement in the House of Commons, 05.09.2011.
- 91 David Cameron, Statement in the House of Commons, 28 February 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110228/debtext/110228-0001.htm#11022819000002 (accessed 5 March 2014).
- 92 Miliband, Statement in the House of Commons, 5 September 2011.
- 93 Jo Swinson, Statement in the House of Commons, 5 September 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110905/debtext/110905-0001.htm#1109054000001 (accessed 4 March 2014).

- 94 David Cameron, Speech on the action we are taking will protect the Libyan people, 18 March 2011. Available from: www.conservatives.com/News/Speeches/2011/03/David_Cameron_Speech_to_Scottish_Conservative_Conference.aspx (accessed 4 March 2014).
- 95 Cameron, Statement in the House of Commons, 5 September 2011.
- 96 Tom Brake, Statement in the House of Commons, 18 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110318/debtext/110318-0001.htm#11031850000007 (accessed 5 March 2014).
- 97 Sam Gyimah, Statement in the House of Commons, 21 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110321/debtext/110321-0001.htm#11032190000002 (accessed 5 March 2014).
- 98 Cameron, Statement in the House of Commons, 21 March 2011.
- 99 William Hague, Statement in the House of Commons, 30 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110330/debtext/110330-0001.htm#11033062000003 (accessed 5 March 2014).
- 100 David Winnick, Statement in the House of Commons, 28 February 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110228/debtext/110228-0001.htm#110228190000002 (accessed 5 March 2014).
- 101 Cameron, Statement in the House of Commons, 18 March 2011.
- 102 Ibid.
- 103 Cameron, Speech on the action we are taking will protect the Libyan people, 18 March 2011.
- 104 Cameron, Statement in the House of Commons, 18 March 2011.
- 105 Hague, Statement in the House of Commons, 30 March 2011.
- 106 Cameron, Statement in the House of Commons, 18 March 2011.
- 107 Cameron, Statement in the House of Commons, 21 March 2011.
- 108 Hague, Statement in the House of Commons, 30 March 2011.
- 109 John McDonnell, Statement in the House of Commons, 18 March 2011. Available from: www.publications.parliament.uk/pa/cm201011/cmhansrd/cm110318/debtext/110318-0001.htm#11031850000007 (accessed 5 March 2014).
- 110 Hague, Statement in the House of Commons, 30 March 2011.
- 111 Cameron, Statement in the House of Commons, 28 February 2011.
- 112 Amnesty International, The Battle for Libya. Killings, Disappearances and Torture, 13 September 2011. Available from: www.amnesty.org/en/library/info/MDE19/025/2011/en (accessed 19 August 2014).
- 113 Ibid., p. 70.
- 114 Human Rights Watch, World Report 2012, Country summary on Libya, p. 598, available at: www.hrw.org/sites/default/files/reports/wr2012.pdf (7 March 2014).
- 115 Ibid., pp. 595, 599.
- 116 Ibid., pp. 598–599.
- 117 Ibid., p. 598.
- 118 Ibid., p. 595.
- 119 Ibid., pp. 598–599.
- 120 Ibid.
- 121 Ibid., p. 599.
- 122 For exceptions see: ‘Rebels tightening grip on Tripoli but the tyrant still eludes them’, *The Times*, 26 August 2011, p. 8.
- 123 Hague, Statement in the House of Commons, 30 March 2011.
- 124 ‘Mercy and justice for Gaddafi’s army dupes’, *The Times*, 21 June 2011, pp. 40–41.
- 125 ‘The fall of Tripoli’, *The Times*, 24 August 2011, pp. 1, 3.
- 126 ‘A Squalid Death’, *The Times*, 26 October 2011, p. 2.
- 127 ‘Arab Spring: Gaddafi’s death: Decaying bodies normally repel – in Misrata, they do exactly the opposite’, *Guardian*, 24 October 2011, p. 3.
- 128 ‘What’s now for Libya?’, *Sun*, 24 October 2011, p. 43.

- 129 'No need to find Gaddafi killer', *Sun*, 25 October 2011, p. 41.
- 130 'National: Death of Gaddafi: In show in a meat locker'. *Guardian*, 22 October 2011, p. 4.
- 131 'Don't you know right from wrong? Irony of Gaddafi's plea to captors', *Sun*, 22 October 2011, pp. 8–9.
- 132 'Gaddafi's killer: "You can have him now"', *The Times*, 22 October 2011, pp. 1, 5.
- 133 'A tyrant meets his end', *The Times*, 21 October 2011, pp. 1, 3.
- 134 'Rot in hell with Hitler', *Sun*, 21 October 2011, pp. 8, 9.
- 135 'Stench of death, flies on blood ... I go inside tyrant's last hideout', *Sun*, 22 October 2011, pp. 8–9.
- 136 Both quotes from 'Rot in hell with Hitler', *Sun*, 21 October 2011, pp. 8–9.
- 137 'National: Death of Gaddafi: Path ahead', *Guardian*, 21 October 2011, p. 2.
- 138 'National: Death of Gaddafi: Legal process: Few Libyans will mourn lost opportunity to see dictator face court', *Guardian*, 21 October 2011, p. 5.
- 139 'Gaddafi's killer: "You can have him now"', *The Times*, 22 October 2011, pp. 1, 5.
- 140 'No need to find Gaddafi killer', *Sun*, 25 October 2011, p. 41.
- 141 Ibid.
- 142 'Arab Spring: Gaddafi's death: Decaying bodies normally repel – in Misrata, they do exactly the opposite', *Guardian*, 24 October 2011, p. 3.
- 143 'A tyrant meets his end', *The Times*, 21 October 2011, pp. 1, 3.
- 144 Ibid.
- 145 'Front: Death of a dictator: Gaddafi killed by rebels in wake of French air strike', *Guardian*, 21 October 2011, p. 1.

References

- Baker, Mona. 2010. 'Narratives of Terrorism and Security: "Accurate" Translations, Suspicious Frames'. *Critical Studies on Terrorism* 3(3): 347–364.
- Brockmeier, Jens, and Donal Carbaugh. 2001. 'Introduction', pp. 1–22 in *Narrative and Identity. Studies in Autobiography, Self and Culture*, y J. Brockmeier and D. Carbaugh, eds. Amsterdam: John Benjamins Publishing Company.
- Browning, Christopher S. 2008. *Constructivism, Narrative and Foreign Policy Analysis. A Case Study of Finland*. Bern: Peter Lang.
- Campbell, David. 1998. *National Deconstruction. Violence, Identity, and Justice in Bosnia*. Minneapolis: University of Minnesota Press.
- Fludernik, Monika. 2009. *An Introduction to Narratology*. London: Routledge.
- Genette, Gérard. 1982. *Figures of Literary Discourse*. New York: Columbia University Press.
- Hansen, Lene. 2006. *Security as Practice: Discourse Analysis and the Bosnian War*. London: Routledge.
- Herman, David. 2002. *Story Logic. Problems and Possibilities of Narrative*. Lincoln: University of Nebraska Press.
- Herman, Luc, and Bart Vervaeck. 2007. 'Ideology', pp. 217–230 in *The Cambridge Companion to Narrative*, D. Herman, ed. Cambridge: Cambridge University Press.
- Herrnstein Smith, Barbara. 1981. 'Narrative Versions, Narrative Theories', pp. 209–232 in *On Narrative*, W.J.T. Mitchell, ed. Chicago, IL: University of Chicago Press.
- Kreiswirth, Martin. 2000. 'Merely Telling Stories? Narrative and Knowledge in the Human Sciences'. *Poetics Today* 21(2): 293–318.
- Kruck, Andreas, and Alexander Spencer. 2013. 'Contested Stories of Commercial Security: Self- and Media Narratives of Private Military and Security Companies'. *Critical Studies on Security* 1(3): 326–346.

- Lacassagne, Aurélie, Tim Nieguth, and François Dépelteau, eds. 2011. *Investigating Shrek. Power, Identity, and Ideology*. Basingstoke: Palgrave.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago, IL: University of Chicago Press.
- Nexon, Daniel H., and Iver B. Neumann, eds. 2006. *Harry Potter and International Relations*. Lanham: Rowman & Littlefield.
- Patterson, Molly, and Kristen Renwick Monroe. 1998. 'Narrative in Political Science'. *Annual Review of Political Science* 1:315–31.
- Phelan, James. 2005. *Living to Tell about It. A Rhetoric and Ethics of Character Narration*. Ithaca: Cornell University Press.
- Ringmar, Erik. 1996. *Identity, Interest and Action: A Cultural Explanation of Sweden's Intervention in the Thirty Year War*. Cambridge: Cambridge University Press.
- Roberts, Geoffrey. 2006. 'History, Theory and the Narrative turn in IR'. *Review of International Studies*, 32(4): 703–714.
- Somers, Margaret. 1994. 'The Narrative Constitution of Identity: A Relational and Network Approach'. *Theory and Society* 23(5): 605–649.
- Spencer, Alexander. 2012. 'The Social Construction of Terrorism. Media, Metaphors and Policy Implications'. *Journal of International Relations and Development* 15(3): 393–419.
- Spencer, Alexander. 2014. 'Romantic Stories of the Pirate in IARRRH: The Failure of Linking Piracy and Terrorism Narratives in Germany'. *International Studies Perspectives* 15(3): 297–312.
- Spencer, Alexander. 2016. *Romantic Narratives in International Politics: Pirates, Rebels and Mercenaries*. Manchester: Manchester University Press.
- Suganami, Hidemi. 2008. 'Narrative Explanation and International Relations: Back to Basics'. *Millennium* 37(2): 327–356.
- Toolan, Michael. 2001. *Narrative: A Critical Linguistic Introduction*. London: Routledge.
- Turner, Mark. 1996. *The Literary Mind*. Oxford: Oxford University Press.

4 PMSCs and Twitter

Sentiment analysis as a tool for evaluating social media data

Magnus Dau and Marlen Martin

Introduction

Private military and security companies (PMSCs) have become a key factor in international security, increasingly performing tasks that were formerly executed by military forces of the state (Singer 2008; Dunigan 2011). The prominence of PMSCs has led scholars to examine the ramifications of their rise (e.g. Avant and Sigelman 2010; Krahmann 2010; Pattison 2014) and the consequences of their work (e.g. Leander 2010; Petersohn 2014). In addition, researchers have studied the self-narratives that PMSCs construct and disseminate on their websites (Joachim and Schneiker 2012, 2014) and the prevailing narratives on PMSCs in traditional news media (Kruck and Spencer 2013). What has been lacking so far is an analysis that compares the dominant frames regarding PMSCs in public discourse outside the traditional media with the frames that PMSCs themselves disseminate to the public. Frames can be understood as a representation of a perspective in which certain aspects are strengthened to influence the generally accepted understanding of a specific image or situation (Joachim 2007). From a liberal perspective, public discourses and the frames they contain help to explain the varying degrees of outsourcing of security services, because public consent to participation in military missions is considered to be influencing the extent of PMSCs' employment (Cusumano 2014: 230–231; Kruck 2014).

In this chapter, we link to this research by studying how PMSCs seek to influence the public via social media and how they are perceived by other actors in these media. Social media platforms such as Twitter, Facebook and YouTube can be considered arenas for public deliberation and the formation of public opinions (Tsaliki 2002; O'Connor *et al.* 2010). The study of the use of social media by business companies is just beginning (e.g. Mangold and Faulds 2009; Hanna *et al.* 2011; Kaplan and Haenlein 2011; Zeitzoff *et al.* 2015). Moreover, scholars of International Relations have relied on social media to analyse a variety of actors, such as terrorist organizations and their self-representations (Farwell 2014); the foreign policies of states (Zetzoff *et al.* 2015); and collective action, as in the case of the Arab Spring (Lotan *et al.* 2011; Hamdy and Gomaa 2012; Wolfsfeld *et al.* 2013). However, the use of social media by PMSCs and social media discourses regarding PMSCs remain to be studied.

Just as PMSCs use their websites (Joachim and Schneiker 2012, 2014), so they presumably also use social media to convey a certain image of themselves. In keeping with the research on PMSCs' self- and media narratives (Kruck and Spencer 2013), we would expect the public discourse on PMSCs in the social media to be more negative than their self-representations. Our aim was to analyse the communication both from PMSCs and about PMSCs in social media. We sought to determine whether there is a difference between these two (sets of) discourses and, if so, to measure the scope of this difference. Another aim was to identify the major frames used by the different sets of actors. Therefore, we focused on the social media platform Twitter and the tweets from the company Academi Elite (formerly known as Blackwater or Xe Services), as well as the tweets about that company. Twitter as a data source is particularly suitable because its users respond promptly to current issues and the limited content of these messages is easy to survey and collect (Zimmer and Proferes 2014: 254ff.).

We chose Blackwater as our subject, because it is one of the best known and most contested PMSCs. Contracted by the US government on a regular basis, the company attracted huge media attention after an incident that occurred during the Iraq War in 2007 (Roberts 2014; Apuzzo 2015), when some Blackwater employees shot 17 innocent Iraqi civilians in Nisour Square in Baghdad.¹ Due to technical limitations imposed by Twitter, we were unable to access tweets that had been published around the time of the Nisour Square incident in 2007. However, we collected other tweets that were issued by Academi Elite about two other events in which the company was involved or that were issued by other Twitter users about these two other events. We began collecting data when rumours emerged that Academi Elite personnel were involved in the crisis in Ukraine and stopped data collection once the trial concerning the Nisour Square massacre had entered its non-public phase. We collected 264,852 tweets that were posted during our data collection period from March through August 2014.

Very large data sets such as ours often pose a challenge when one attempts to analyse social media (Lewis *et al.* 2013). Ultimately, we were interested in a qualitative aspect; that is, the frames used by Academi and by others concerning Academi. However, analysing this enormous amount of data by immediately using qualitative methods (e.g. qualitative content analysis) was not feasible because of the inherently time-consuming nature of such methods and the huge dataset. Instead, we decided to begin with a quantitative method that would allow us to obtain an overview of the data and to categorize it and thereby to reduce the dataset that we then could analyse with qualitative means; we opted for a sentiment analysis.

Sentiment analysis, which is also referred to as opinion mining, is a computational technique to automatically categorize text segments according to their evaluative expressions (Pang and Lee 2008; Liu 2012, 2015; Lucas *et al.* 2015; Serrano-Guerrero *et al.* 2015). Bing Liu has described sentiment analysis as the 'field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes' (Liu 2012: 7). 'Sentiment

analysis’ is commonly used as an umbrella term to denote various techniques and research interests in the field of natural language processing (NLP) that seek to assess textual representations of human communication quantitatively. All these techniques have in common the evaluation of source material to identify its general sentiment in terms of its positive or negative polarity on a predefined scale.

For our sentiment analysis of social media discourses on PMSCs, we used a software program based on the Stanford Sentiment framework that classifies tweets into five categories of evaluative statements: very positive, positive, neutral, negative and very negative. Such categorization alone does not allow us to determine which frames are disseminated in the social media, but it does enable us to identify different sets of tweets that can then be analysed by qualitative means. In our case, the sentiment analysis allowed us to identify the very positive and the very negative tweets, which we then subjected to qualitative content analysis. More generally, applying sentiment analysis to identify relevant sets of data drawn from a very large database can be useful, for example, in evaluating how political actors are perceived in social media (Tumasjan *et al.* 2010; Murthy 2015); in measuring opinions regarding a particular policy (Ceron *et al.* 2013; González-Bailón and Paltoglou 2015), company or institution (Van de Kauter *et al.* 2015); or for studying how social media are used by terrorists (Cheong and Lee 2011; Klausen 2015).

According to Bryman’s (2006) typology of mixed-methods designs, our methodological approach represents an ‘offset’ design rationale, because both quantitative sentiment analysis and qualitative content analysis are employed to mitigate the weaknesses of each of these instruments and to combine their strengths. The strength of the sentiment analysis is its power to categorize large amounts of data. This type of computer-assisted content analysis is ‘well-suited to identifying and confirming “blank spots” in discourse, or revealing what is not talked about or what is no longer talked about’ (Bennett 2015: 997). By applying qualitative content analysis to specific categories of tweets identified in the sentiment analysis, we were able to reconstruct and interpret the dominant positive and negative frames with respect to Academi that appeared in the discourse on Twitter (see also Bryman 2012).

The following sections explain the steps necessary in collecting and preparing the data and subsequently conducting a sentiment analysis. The chapter concludes with an overview of our findings, with a special focus on the sets of statements we identified as being the most negative and the most positive.

Data retrieval and preparation

Retrieving social media content on a large scale cannot be done manually. Opening the Twitter website and manually extracting tweets, for example, is a slow and error-prone process. Commercial businesses offer tailored data sets with prices ranging from several hundred to thousands of US dollars, usually depending on the volume. Another way of obtaining a complete data set is to use special software tools that can retrieve data to meet researchers’ specific needs

by filtering the data according to keywords, geographic areas or languages. In addition to a wide range of fully or partially commercial tools, there are also free and open-source solutions such as NodeXL and the package *twitterR* for the R statistical computing project.

Our data sets were collected by means of a program we wrote to access the Twitter API.² This program allowed us to collect all tweets originating from and directed to Academi, as well as all tweets that contained the keywords ‘Academi Elite’, ‘Blackwater’ or ‘Xe Services’ that originated from other users. This enabled us to collect 264,852 tweets, which were stored in ordinary text documents of medium size.³ We then divided the data into two data sets to identify frames that were dominant in the public and the company spheres. The first data set contained the tweets published by Academi Elite and those directly addressed at it using, for example, an ‘@’ reference; this will now be referred to as the ‘Academi data set’. The second data set of tweets, which will be referred to as the ‘public data set’, consisted of all tweets that did not originate from the company itself but that contained the current name of Academi Elite or one of its former names (e.g. ‘Blackwater’). The two data sets were not equal in size: the Academi data set contained 1356 tweets, whereas the public data set consisted of 263,496 tweets.

The next step was to consolidate the data sets by removing all tweets that were not relevant to our research question. Owing to the large size of the resulting data set, we extracted a random sample of 1000 tweets, which we then processed manually to identify and remove all non-fitting tweets along with their respective retweets from the overall data set. A significant portion of the public data set contained tweets that were not related to our case study of the PMSC Academi Elite but referred to the popular TV series *Game of Thrones*, the promotion of songs, or a women’s football match, among many other things.⁴ In addition, all non-English tweets were removed from the data sets, because the aforementioned software program works best when the English language is used. After the removal of these tweets, the public data set contained a total of 162,350 entries and the Academi data set remained the same. We decided not to filter out tweets with identical text, because they occurred in the form of retweets and expressed the implicit support and sharing of the original message.

To ensure smooth processing by the sentiment analysis program, the collected data first had to be revised. Text contained in tweets differs significantly from most other types of written language, such as that seen in newspaper articles, books or blog entries. In addition to text messages, tweets may contain links to other websites or may address another Twitter user directly by their user name in the format ‘@username’. Such particularities typically cause problems, because most natural language processing techniques, such as sentiment analysis, work best with or even require correct formal language. In contrast, tweets often have a complex structure, are idiosyncratic and usually contain colloquial language. These properties make them a tough challenge for linguists and computer scientists alike who are trying to improve the automated evaluation of natural language. One example of such difficulties is the seemingly easy task of identifying

the end of a sentence. Although in most cases sentences are separated by a period, question mark or exclamation mark, a multitude of exceptions to this rule exist with respect to tweets. Not only are some entities often referred to by their abbreviations (e.g. ‘U.S.’); certain titles indicating a level of professional advancement such as ‘Dr.’ or quantities such as ‘4 million umbrellas’ require a sophisticated approach to sentence detection.

In addition, computer programs still struggle with certain characteristics of human language, such as sarcasm, irony, allusions, innuendo, dialects and metaphors. For instance, the colloquial sentence ‘This was a funny movie – not!’ might be interpreted by the program as a very positive statement because of its grammatical structure. For the reader, however, the unconventional form in which the negation appears at the end of the sentence changes the meaning entirely and is reliably and correctly interpreted as a negative statement.

Other popular characteristics of Twitter, such as hashtags, further complicated processing for us. Hashtags must be preceded by a hash sign (#), and spaces between its components are not allowed, which leads to the ‘gluing together’ of words or phrases (e.g. in the hashtag #ukrainiancrisis). The multitude of slightly different hashtags referring to the same event further aggravates the problem, and to optimally use the maximum of 140 characters allowed in a tweet, users often rely on abbreviations or invent new, creative expressions such as ‘w/o’ (‘without’) and ‘w/e’ (‘whatever’). As the technical development of the underlying technologies progresses, entirely new forms of textual representations are emerging on Twitter, including smileys (‘:-’) and the more diverse emojis. The latter are graphical representations frequently used in social media; for example, the ‘thumb-up’ icon is derived from Twitter.

To bridge the gap between the expected formal language and the language that is actually found on Twitter, we removed all links and the expression ‘RT’ that are automatically added to a tweet when it is retweeted. We also replaced user references with invented surnames, such as ‘Daniel’, and expanded the most commonly used abbreviations (e.g. replacing ‘ive’ with ‘I have’). A common phenomenon that was more difficult to address was the expansion of hashtags, such as ‘#unitedstatesdefenseandmilitaryforces’; still, a significant proportion of the hashtags could be processed by simply removing the #. In the case of ‘#Iraq’, for example, we decided to remove only the preceding # from the tweets. After preparing the tweets of both data sets in this way, we applied the sentiment analysis.

Sentiment analysis

The most widespread approaches to sentiment analysis rely on machine-learning techniques in which algorithms are trained on small samples of data until they have reached a satisfactory level of accuracy. Machine-learning techniques have become widely used because they usually offer high levels of accuracy and can be applied in a wide range of scenarios, even if the source material poorly adheres to syntactic or semantic conventions, as is the case with tweet content. A

key feature of such machine-learning frameworks and algorithms is their ability to construct patterns from a limited amount of data and to then apply these patterns to new data. In most cases, machine-learning techniques yield very good results, but creating them can be time-consuming.⁵ To conduct the sentiment analysis in our research project, we used the well-documented open-source framework Stanford NLP from the Stanford Natural Language Processing Group.⁶

Being a machine-learning framework, the Stanford program requires a model to analyse the data. The model provides a benchmark for the computer program to sort the sentiments into the proper categories, and the process of building a model is called training. The standard Stanford Sentiment framework was originally trained on 11,855 sentences derived from online film reviews, each sentence of which was annotated with a score according to a five-level scale ranging from 0 to 4: very negative (0), negative or fairly negative (1), neutral (2) positive or fairly positive (3) and very positive (4). For example,

Original text: ‘Good film, but very glum.’

Annotated version: (3 (4 (3 Good) (2 film)) (2 (2,) (1 (2 (2 but) (2 very)) (1 (1 glum) (2.)))))⁷

The numbers preceding the words in parentheses indicate the sentiment scores for the individual word or phrase, and the words are then grouped together with additional parentheses to indicate their relation. Next, the grouped words are assigned a new sentiment score to account for changes that might result from the regrouping. In the example shown above, for instance, the word ‘good’ was assigned a positive score (3) and the word ‘film’ a neutral score (2), but together the words express a very positive sentiment, so as a group they were assigned the highest score for positivity (4).

Sentiment scores are also assigned on the level of phrases. Each phrase is extracted from the tweet’s text and in most cases represents a complete sentence. Table 4.1 shows the distribution of recognized phrases per tweet, indicating that most tweets consist of no more than two phrases (or sentences).

Table 4.1 Distribution of sentences per tweet in the data sets

<i>Sentences per tweet</i>	<i>Academi Elite data set (%)</i>	<i>Public data set (%)</i>
1	42	64
2	47	18
3	13	13
4	1	4

Source: own table.

Note

Rounding errors may lead to sums other than 100%. Rows representing less than 1% were omitted.

If a tweet consists of more than one phrase, how should the final score for the whole tweet be calculated? One way to do this would be to determine the average of all the values. Another would be to take only the score of the longest phrase in a *pars pro toto* manner, based on the argument that the longest phrase carries more meaning because it is more elaborate and therefore more indicative of the sentiment than are the other phrases. However, this need not be the case, as can be seen in the following example:

‘No one could do anything about it as we were in Iraq.’ [phrase 1] (2)
 ‘– ISIS on gains?’ [phrase 2] (2)
 ‘Nope, Blackwater on killing US govt workers.’ [phrase 3] (1)
 ‘<http://t.co/3VL94SezKy>’ [phrase 4] (2)⁸

The sentiment analysis generates a neutral score (2) for the first, second and fourth phrases, whereas a fairly negative score (1) was assigned to the third phrase.⁹ Although the first phrase is the longest part of the tweet, only the third phrase clearly indicates the negative sentiment of the tweet, so the overall sentiment would be neutral. The arithmetic mean of 1.75 (7 divided by 4) still indicates a largely neutral tweet, with a weighted average of 1.63. The example shows how the addition of a hyperlink in the tweet may blur the result in that it distorts the calculated results of the sentiment analysis. By removing the link, we can reduce the number of phrases to three and the arithmetic mean score to 1.67 (5 divided by 3) and the weighted average to 1.57. For this reason and in order to increase the precision of our results, all links in the tweets were removed.

In our research project, the question of what strategy should be used to compute the final score for each text was actually more of a thought experiment. Owing to the 140-character length limitation, the tweets in our study consisted mostly of one sentence, rarely of two and very rarely of three or more. A pretest had revealed that the different calculating schemes produced similar outputs for our data, so we could have used them interchangeably in our example; we decided to base our results on the scores assigned to the longest text phrase detected in a given tweet.

Because the researchers at Stanford University had trained their program on 11,855 sentences obtained from online film reviews, it had to be retrained for the coding of tweets. We manually coded 1000 tweets out of the Academi data set and the public data set by assigning sentiment scores in a syntactically and semantically plausible manner (see Socher *et al.* 2013). The (re-)training of the model required three data sets: the training data set, the development data set and the test data set, each of which represented a different fraction of the data under scrutiny. First, we determined the sentiment scores for 1000 tweets based on manual coding, after which we divided the 1000 manually coded tweets into the three data sets to train the program on how to classify tweets.

The first data set (the training data set) was used to train a model of the program, which was then tested on the second data set (the development data set). In an iterative procedure, in which the model was trained and tested on the

development model several times, the model's accuracy was gradually improved. After the training, the model's accuracy was assessed by applying it to a third data set, which by convention is called the test data set. When the accuracy of the model reached 70.12 per cent, we considered it to be sufficiently trained.¹⁰

Results of the sentiment analysis

Our expectations regarding the results of the sentiment analysis were based on two hypotheses: (1) that the sentiment scores for the tweets in the Academi data set would differ from those in the public data set and (2) that the company's corporate image as disseminated on Twitter was likely to be presented in neutral or even positive tones, whereas the tweets representing public discussion involving the company were more likely to be negative owing to the scandals in which Blackwater was involved. Figure 4.1 presents the results of the sentiment analysis for both data sets. The neutral and moderate categories contained the vast majority of tweets, whereas tweets with very positive or outright negative scores were fairly rare. Thus, our sentiment analysis confirmed both our first hypothesis (i.e. that the sentiment scores would differ across the two data sets) and our second hypothesis (i.e. that the Academi data set would generally yield more positive sentiment scores than the public data set would).

In accordance with our mixed-methods approach, the next step after the sentiment analysis was to select the very positive and very negative tweets in each data set for further qualitative content analysis. For the Academi data set, the scoring revealed 15 very negative tweets and 34 very positive tweets; in the public data set, there were 9032 very negative tweets and 813 very positive tweets. As expected, the tweets from the public data set contained generally more negative notions when compared with the Academi data set; nonetheless,

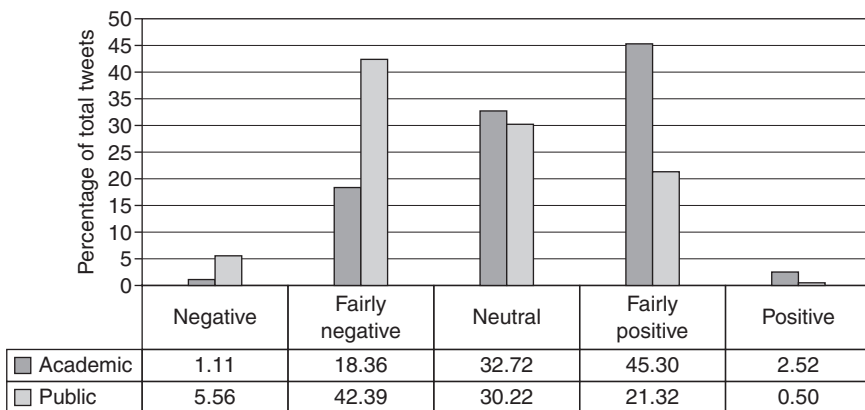


Figure 4.1 Comparison of sentiment scores as percentages of all tweets in the data sets.

Source: own figure.

the share of negative tweets in the Academi data set was significant and worth looking into.

Qualitative content analysis of outlier values

A closer qualitative examination of the tweets was necessary to understand which elements of the tweets were decisive for the categorization ‘very negative’ or ‘very positive’. We saved all the negative tweets in a text file and qualitatively examined a random sample of 1000 entries in a *pars pro toto* manner. We found that the prevalence of negative tweets can at least partially be attributed to the specific topics surrounding PMSCs and the related vocabulary. As a PMSC, Academi publishes security-related news that is evaluated as negative by the program used to classify the tweets for the sentiment analysis. To give an example, the following tweet, posted by Academi, was scored as fairly negative because it contains the words ‘killed’ and ‘attack’:

Policeman killed in al-Shabaab attack in Kenya. Read more from India Times: <http://t.co/knE5YPMefK> #news #kenya #al-shebaab.

Although the content of the information is evidently negative, it is unlikely that it would have a negative tone considering that it is Academi’s own news tweet. In particular, tweets related to the company’s work, such as workshops on gun safety (e.g. ‘We have a variety of training options for civilians and individuals: pistols, shotguns, rifles – we have it all!’) result in negative sentiment scores, probably because the tweet refers to specific weapons. However, this does not, or not necessarily, lead to a negative evaluation of the PMSC; rather, in the case of a PMSC, words that would otherwise connote negativity might not carry the same meaning. This characteristic represents a challenge to research on PMSCs or on armed security actors more generally.¹¹ Thus, our qualitative analysis revealed problems with the sentiment analysis that would otherwise have remained unnoticed.

Moreover, our qualitative content analysis (Mayring 2008) yielded insights into the prevalent topics and arguments that underlie the positive or negative evaluations of Academi on Twitter. We conducted separate qualitative content analyses of the Academi data set and the public data set, focusing on the manifest content of tweets categorized as very negative or very positive. The messages were then summarized based on the most common topics.

Two topics dominated the public discussions about Academi: the threatening of a state department official by an employee of the company (then operating under the name Blackwater) and the trial following the shooting in Nisour Square in Baghdad. Both incidents occurred in 2007, at which time the company was a major contractor working for the US government in Iraq.

The first event, though it dated back to 2007, gained momentum in the Twitter community only later, in 2014, when an article published by the *New York Times* revealed that a state department official had been threatened by a Blackwater

employee (see Risen 2014). The incident as a whole, especially the role of the company and its employees, was received very negatively, as indicated by the sentiment scores. This finding was substantiated by a qualitative evaluation of tweets condemning the incident, such as '[t]his is insane report: blackwater threatened to kill a state department official investigating them'.

Even more prominent than the first event, and received almost as negatively, was the trial that led to the conviction of former Blackwater employees for the shooting in Baghdad's Nisour Square in 2007 in which several Iraqi civilians were killed. The trial took place in a federal court in Washington, DC. The US Department of Justice charged the former Blackwater operatives with counts of first-degree murder. Right from the beginning the trial was discussed so intensively that the judge, Royce Lamberth, explicitly ordered the jurors not to consult any social media at all – specifically mentioning Twitter – but to stick to the evidence presented during the trial, reportedly remarking, 'Just don't do it; just stick to the evidence' (Al Jazeera America 2014).

Several tweets indicate that Twitter users approved of the Blackwater guards having to face charges: 'Blackwater operatives finally being tried for murders? Am I dreaming? <http://t.co/znNripwOuv>' and '[...]. This is awesome. Next Blackwater Academi should be sued to Hell and back: #Mercenaries should be #hostishumanigener...'. Some tweets suggested incomprehension and indignation about the continuous contracting of the company even after the 2007 incident. For economic reasons, some even questioned the need to hire PMSCs in general. Users felt that Blackwater was paid too much for duties that the US military could and should have performed itself, as indicated by the tweet 'Too much spent on military? Probably. Too much spent on outsourced privateers of the Blackwater ilk? FFS yes!'. Moreover, users expressed the fear that the government did not have enough control over the contractors or that those companies themselves had too much power. This is evident in tweets such as 'Holy shit, lost control over your mercenaries, US? Blackwater manager threatened to kill govt inspector ~ <http://t.co/eQXM6YBWYk>' or 'Are contractors too powerful?@nytimes article on Blackwater running the show in Iraq makes me think of Iron Man: <http://t.co/4CT8oUMcpL>'. The issue of privatizing security was rarely mentioned, although the debate clarifies that some users considered PMSCs per se as being problematic: 'There are a lot of government functions that should never be privatized. Defense is one of those. #Blackwater.'

In addition to the tweets condemning the incident in Nisour Square and the subsequent trial of those involved, as well as the tweets criticizing Blackwater, statements also appeared that seemingly came from former operatives of Blackwater and that criticized the whole trial with comments such as 'Its disgusting and wrong and i didnt serve 3 tours as a blackwater/xe mercenary to come home to this bullshit.'

The aforementioned two topics (the event and the trial) were both derived from the strongly negative tweets of the public data set. But, perhaps more surprisingly, criticism of the company could also be found in the tweets that contained very positive sentiments. Despite their positive expression, these tweets

shared the criticism of the negative tweets but did so in a syntactically and semantically positive tone. For example, one tweet strongly lauded investigative journalist Jeremy Scahill's book in which he strongly criticizes Blackwater: '[h]ighly recommend jeremy scahill's blackwater: the rise of the worlds most powerful mercenary army. it's a great read.' Thus, the qualitative content analysis revealed that both the negative and the positive tweets in the public data set were largely dominated by the same strong criticism of Blackwater and differed only in their articulation. This shows that the qualitative content analysis was crucial to complement (and at times correct) the preceding sentiment analysis.

The topics that dominated the public data set were virtually absent from the Academi data set, which consisted primarily of company PR messages to promote training courses and job openings. The most frequently occurring topic in the positive tweets were references to the 100,000 job initiatives designed to provide jobs to veterans and Academi's contribution to this initiative, such as 'Academi congratulates the 100k jobs mission and is proud to contribute to its goal!'

Again, the limits of the sentiment analysis are shown in the case of tweets from the Academi data set that received negative scores. Tweets such as 'Record number of U.N. peacekeepers fails to stop african wars from [...]', which was linked to an article in the *Washington Post*, were classified as very negative by the program, even though the message as relayed by Academi was little more than a fairly neutral newspaper article and did not indicate a negative assessment of Academi. The negative score was a result of the combination of the idea of failing to stop wars and the word 'war' itself. Nonetheless, the sentence as a whole does not indicate a negative tone but rather addresses a topic that can easily be interpreted as carrying a negative meaning. The negatively classified tweets of the Academi data set consisted almost exclusively of such false positives. Correcting these wrong classifications required manual recoding of the corresponding tweets and adding them to the material on which the sentiment model was trained.

Conclusion

The goal of our analysis was (1) to test whether differences in the tonal polarity were evident between the discourses of the PMSC Academi Elite and those of the general public on Twitter about Academi Elite, and (2) to identify the crucial frames responsible for different tonalities. The amount and linguistic characteristics of the data posed serious methodological challenges that were overcome by means of a mixed-methods design that combined quantitative sentiment analysis and qualitative content analysis. Using this design, we relied both on the sentiment analysis to provide evidence of differences between the two data sets with regard to the expressed sentiments and on the content analysis to determine that the dominant frames were significantly more negative in the public discourses about the PMSC than they were in the company's immediate Twitter sphere.

Classifying the tweets by means of the sentiment analysis enabled the subsequent content analysis. The qualitative content analysis showed that there were no topical intersections between the two sets of discourses and identified the key frames responsible for the different sentiment scores for each data set. It also largely confirmed the differences in the distribution of sentiment scores indicated by the sentiment analysis and permitted erroneous or misleading assessments of the sentiment analysis to be corrected.

Our study provides several insights into the use of sentiment analysis. Its greatest advantage is its strength in processing large amounts of data in an era when massive data sets are becoming the norm for researchers in the social sciences. As a heuristic tool, it gauges the distribution of sentiments within data sets, enabling us not only to compare them with one another but also to perform further analysis by selecting the most promising data for subsequent, in-depth (qualitative) analysis. Our content analysis not only corroborated our initial hypotheses – which had already been confirmed by the results of the sentiment analysis – but also allowed us to identify the major arguments and thematic aspects present in each data set, which led to either a negative or a positive representation of the company.

Notes

- 1 However, we are aware of the singularity of Blackwater and would like to emphasize that it is not representative of the PMSC industry – most PMSCs do not engage in armed activities, whether in combat or otherwise.
- 2 API is an abbreviation for application programming interface.
- 3 The individual document size did not exceed 30 MB.
- 4 The identified tweets and their retweets were removed from the sample.
- 5 For an alternative approach, such as dictionary-based techniques, see Grimmer and Stewart (2013); Lucas *et al.* (2015); Serrano-Guerrero *et al.* (2015).
- 6 The framework is available online at <http://nlp.stanford.edu>. Alternatives include the Apache Software Foundation Mahout (<http://mahout.apache.org>) and its Spark framework (<http://spark.apache.org/mllib>).
- 7 A training, development and testing data set is available from <http://nlp.stanford.edu/sentiment>.
- 8 We anonymized all the tweets cited in this chapter to protect the privacy of the Twitter users.
- 9 As indicated by the scores for the links, the software framework assigns scores to phrases even if they are not part of a strictly syntactically and semantically correct sentence. Even if these scores were always neutral – which they are not – they would distort the result of the calculation of the score for the whole sentence. Therefore, we removed all links in the tweets prior to processing.
- 10 There is no universally accepted threshold for the level of accuracy; accuracy is understood here as the portion of correct classifications of the overall sentiment of a phrase. The model trained here reached a level of accuracy of 70.12 per cent. In comparison, machine-learning programs have achieved rates ranging from 63.5 per cent (Diakopoulos *et al.* 2011: 196ff.) to 80 per cent or better (Annett and Kondrak 2008), whereas trained human coders usually exhibit an intercoder reliability of 70 to 80 per cent (Mikhaylov *et al.* 2012: 83). Therefore, we considered an accuracy level of about 70 per cent to be sufficient for our purposes.

- 11 To mitigate the impact of this phenomenon, the sentiment model could be retrained using tweets that exhibit this pattern. We realized this only after our first round of analyses; it shows that sentiment analysis based on machine-learning is an iterative process. For optimal results, the model would have to be retrained as suggested. Given the low number of tweets exhibiting this pattern and its consequently negligible impact, we did not pursue this option.

References

- Al Jazeera America. 2014. 'Blackwater Guards Face Trial in Iraq Shootings'. *Al Jazeera America*, 11 June.
- Annett, Michelle, and Grzegorz Kondrak. 2008. 'A Comparison of Sentiment Analysis Techniques: Polarizing Movie Blogs', pp. 25–35 in *Advances in Artificial Intelligence: Proceedings of the 21st Conference of the Canadian Society for Computational Studies of Intelligence, Canadian AI 2008, Windsor, Canada, May 28–30, 2008*, R. Goebel, J. Siekmann, and W. Wahlster, eds. Berlin: Springer.
- Apuzzo, Matt. 2015. 'Ex-Blackwater Guards Given Long Terms for Killing Iraqis'. *New York Times*, 13 April.
- Avant, Deborah, and Lee Sigelman. 2010. 'Private Security and Democracy: Lessons from the US in Iraq'. *Security Studies* 19(2): 230–265.
- Bennett, Andrew. 2015. 'Found in Translation: Combining Discourse Analysis with Computer Assisted Content Analysis'. *Millennium: Journal of International Studies* 43(3): 984–997.
- Bryman, Alan. 2006. 'Integrating Quantitative and Qualitative Research: How Is It Done?' *Qualitative Research* 6(1): 97–113.
- Bryman, Alan. 2012. *Social Research Methods*. 4th edn. Oxford: Oxford University Press.
- Ceron, Andrea, Luigi Curini, Stefano M. Iacus, and Giuseppe Porro. 2013. 'Every Tweet Counts? How Sentiment Analysis of Social Media Can Improve Our Knowledge of Citizens' Political Preferences with an Application to Italy and France'. *New Media & Society*. doi:10.1177/1461444813480466.
- Cheong, Marc, and Vincent C. S. Lee. 2011. 'A Microblogging-Based Approach to Terrorism Informatics: Exploration and Chronicling Civilian Sentiments and Response to Terrorism Events via Twitter'. *Information Systems Frontiers* 13(1): 45–59.
- Cusumano, Eugenio. 2014. 'The Scope of Military Privatisation: Military Role Conceptions and Contractor Support in the United States and the United Kingdom'. *International Relations*. doi:10.1177/0047117814552142.
- Diakopoulos, Nicholas, Mor Naaman, Tayebbeh Yazdani, and Funda Kivran-Swaine. 2011. 'Social Media Visual Analytics for Events', pp. 189–212 in *Social Media Modeling and Computing*, S.C.H. Hoi, J. Luo, S. Boll, D. Xu, R. Jin, and I. King, eds. London: Springer.
- Dunigan, Molly. 2011. *Victory for Hire: Private Security Companies' Impact on Military Effectiveness*. Stanford: Stanford University Press.
- Farwell, James P. 2014. 'The Media Strategy of ISIS'. *Survival: Global Politics and Strategy* 56(6): 49–55.
- González-Bailón, Sandra, and Georgios Paltoglou. 2015. 'Signals of Public Opinion in Online Communication: A Comparison of Methods and Data Sources'. *The ANNALS of the American Academy of Political and Social Science* 659(1): 95–107.
- Grimmer, Justin, and Brandon M. Stewart. 2013. 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts'. *Political Analysis* 21(3): 267–297.

- Hamdy, Naila, and Ehab H. Gomaa. 2012. 'Framing the Egyptian Uprising in Arabic Language Newspapers and Social Media'. *Journal of Communication* 62(2): 195–211.
- Hanna, Richard, Andrew Rohm, and Victoria L. Crittenden. 2011. 'We're All Connected: The Power of the Social Media Ecosystem'. *Business Horizons* 54(3): 265–273.
- Joachim, Jutta M. 2007. *Agenda Setting, the UN, and NGOs: Gender Violence and Reproductive Rights*. Washington, DC: Georgetown University Press.
- Joachim, Jutta, and Andrea Schneiker. 2012. 'Of "True Professionals" and "Ethical Hero Warriors": A Gender-Discourse Analysis of Private Military and Security Companies'. *Security Dialogue* 43(6): 495–512.
- Joachim, Jutta, and Andrea Schneiker. 2014. 'All for One and One in All: Private Military Security Companies as Soldiers, Business Managers and Humanitarians'. *Cambridge Review of International Affairs* 27(2): 246–267.
- Kaplan, Andreas M., and Michael Haenlein. 2011. 'Two Hearts in Three-Quarter Time: How to Waltz the Social Media/Viral Marketing Dance'. *Business Horizons* 54(3): 253–263.
- Klausen, Jytte. 2015. 'Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq'. *Studies in Conflict & Terrorism* 38(1): 1–22.
- Krahmann, Elke. 2010. *States, Citizens and the Privatization of Security*. Cambridge, UK: Cambridge University Press.
- Kruck, Andreas. 2014. 'Theorising the Use of Private Military and Security Companies: A Synthetic Perspective'. *Journal of International Relations and Development* 17: 112–141.
- Kruck, Andreas, and Alexander Spencer. 2013. 'Contested Stories of Commercial Security: Self- and Media Narratives of Private Military and Security Companies'. *Critical Studies on Security* 1(3): 326–346.
- Leander, Anna. 2010. "The Paradoxical Impunity of Private Military Companies: Authority and the Limits to Legal Accountability." *Security Dialogue* 41(5):467–90.
- Lewis, Seth C., Rodrigo Zamith, and Alfred Hermida. 2013. 'Content Analysis in an Era of Big Data: A Hybrid Approach to Computational and Manual Methods'. *Journal of Broadcasting & Electronic Media* 57(1): 34–52.
- Liu, Bing. 2012. *Sentiment Analysis and Opinion Mining*. San Rafael: Morgan & Claypool.
- Liu, Bing. 2015. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge, UK: Cambridge University Press.
- Lotan, Gilad, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and Danah Boyd. 2011. 'The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions'. *International Journal of Communication* 5: Feature 1375–1405.
- Lucas, Christopher, Richard Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. 'Computer Assisted Text Analysis for Comparative Politics'. *Political Analysis* 23(2): 254–277.
- Mangold, W. Glynn, and David J. Faulds. 2009. 'Social Media: The New Hybrid Element of the Promotion Mix'. *Business Horizons* 52(4): 357–365.
- Mayring, Philipp. 2008. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 2nd edn. Weinheim: Beltz.
- Mikhaylov, Slava, Michael Laver, and Kenneth R. Benoit. 2012. 'Coder Reliability and Misclassification in the Human Coding of Party Manifestos'. *Political Analysis* 20(1): 78–91.
- Murthy, Dhiraj. 2015. 'Twitter and Elections: Are Tweets Predictive, Reactive, or a Form of Buzz?' *Information, Communication & Society* 18(7): 816–831.

- O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. 2010. 'From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series', pp. 122–129 in *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, May 23–26, 2010, Washington, DC*. Menlo Park: Association for the Advancement of Artificial Intelligence.
- Pang, Bo, and Lillian Lee. 2008. 'Opinion Mining and Sentiment Analysis'. *Foundations and Trends in Information Retrieval* 2(1/2): 1–135.
- Pattison, James. 2014. *The Morality of Private War: The Challenge of Private Military and Security Companies*. Oxford: Oxford University Press.
- Petersohn, Ulrich. 2014. 'The Impact of Mercenaries and Private Military and Security Companies on Civil War Severity between 1946 and 2002'. *International Interactions: Empirical and Theoretical Research in International Relations* 40(2): 191–215.
- Risen, James. 2014. 'Before Shooting in Iraq, a Warning on Blackwater'. *New York Times*, 29 June.
- Roberts, Dan. 2014. 'US Jury Convicts Blackwater Guards in 2007 Killing of Iraqi Civilians'. *Guardian*, 23 October.
- Serrano-Guerrero, Jesus, José A. Olivas, Francisco P. Romero, and Enrique Herrera-Viedma. 2015. 'Sentiment Analysis: A Review and Comparative Analysis of Web Services'. *Information Sciences* 311: 18–38.
- Singer, Peter Warren. 2008. *Corporate Warriors: The Rise of the Privatized Military Industry*. Ithaca: Cornell University Press.
- Socher, Richard, Alex Perelygin, Jean Y. Wu, Jason Chuang, Christopher D. Manning, Andrew Y. Ng, and Christopher Potts. 2013. *Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank*. Stanford: Stanford University.
- Tsaliki, Liza. 2002. 'Online Forums and the Enlargement of Public Space: Research Findings from a European Project'. *Javnost – The Public: Journal of the European Institute for Communication and Culture* 9(2): 95–112.
- Tumasjan, Andranik, Timm O. Sprenger, Philipp G. Sandner, and Isabell M. Welp. 2010. *Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment*. Munich: Technical University of Munich.
- Van de Kauter, Marjan, Diane Breesch, and Véronique Hoste. 2015. 'Fine-Grained Analysis of Explicit and Implicit Sentiment in Financial News Articles'. *Expert Systems with Applications* 42(11): 4999–5010.
- Wolfsfeld, Gadi, Elad Segev, and Tamir Sheafer. 2013. 'Social Media and the Arab Spring: Politics Comes First'. *The International Journal of Press/Politics*. doi:10.1177/1940161212471716.
- Zeitsoff, Thomas, John Kelly, and Gilad Lotan. 2015. 'Using Social Media to Measure Foreign Policy Dynamics: An Empirical Analysis of the Iranian–Israeli Confrontation (2012–13)'. *Journal of Peace Research*. doi:10.1177/0022343314558700.
- Zimmer, Michael, and Nicholas J. Proferes. 2014. 'A Topology of Twitter Research: Disciplines, Methods, and Ethics'. *Aslib Journal of Information* 66(3): 250–261.

5 Semi-structured interviews with non-state and security actors

Anja Mihr

Introduction

In this chapter, I will discuss how to use semi-structured interviews with non-state security actors, based on my experience and expertise gathered during different empirical research projects. An interview is a method to collect data in which one person, the interviewer, asks questions to another person, the respondent. This can be done either face to face, by telephone or Skype or by email. In the social sciences, interviews are among the most common methods used to collect data (DiCicco-Bloom and Crabtree 2006; Polit and Beck 2006). On the one hand, interviews are based on in-depth preparation prior to the interview. On the other hand, they rely on careful analytical assessment of the interview afterwards as well as on contextual analysis.

An interview may be either structured or semi-structured (Fontana and Frey 2005). Structured interviews use a questionnaire format that includes specific (closed) questions and ‘boxes to be ticked’. In most cases, they are used when language translation is needed, the respondent has a language impairment or the interview is being conducted by email. Structured interviews are used to generate quantitative rather than qualitative data. Semi-structured interviews involve an ‘interview guide’ rather than a questionnaire; that is, a set of questions that lead the interviewer through the session with the respondent. There are no boxes to tick, and the data generated are of a qualitative nature. Semi-structured interviews allow for a reflexive approach towards the respondent; that is, interviewers may ask respondents more tailored questions and engage them in a conversation about the subject matter at hand. If time allows, both sides have the flexibility to delve deeper into a particular topic.

Semi-structured interviews are most useful when researchers have only one chance to interview someone and when several interviewers and researchers of the same project will be sent out to do the fieldwork at different stages and in different countries and with different interviewees. According to Bernard (2006), the semi-structured interview guide provides a clear set of instructions for interviewers and helps produce comparable qualitative data. The gathered data allow for interpretation to answer the specific research questions of the project, program or fieldwork. Semi-structured interviews are often preceded or accompanied by participatory

observations (see Chapter 12 by Tessa Diphoom) or informal and unstructured talks, such as in conjunction with a political event, caucus or conference. This may help the researcher gain a keen understanding of the topic and develop further semi-structured questions for subsequent interviews with specific key stakeholders that are relevant to the project.

Semi-structured interview guides allow researchers to ask open-ended questions and to interpret the ‘why’ question of a study in many research contexts, whereas structured interviews tend to answer ‘how many’ questions and are more suitable for data collection, surveys (see Chapter 13 by Sabrina Karim) and statistics. The difference is of particular importance in inter- and transdisciplinary research, which involves researchers from different backgrounds who conduct interviews at different stages or phases of a project. Before a research project is begun, those involved in its design should make clear to the project team which type or types of interview(s) will be used, even if they decide to combine structured and semi-structured approaches. In addition, there is the question of whom to interview, which in the case of non-state security actors can present particular challenges.

An interview guide is a useful tool for interviewing non-state security actors, whether they be members of the local paramilitary or of a private, commercial security sector training company, such as Academi (formerly Blackwater). These actors often emerge when governments and local authorities are unable to provide effective and reliable security and protection. In many countries in transition, after times of conflict or war, these non-state actors are usually one of the primary targets for security sector reform (SSR). But, more generally speaking, non-state actors in the field of security are diverse and broad-ranging; they include non-governmental organizations (NGOs), civil society organizations (CSOs), multinational or national corporations, religious groups, and diaspora, migrant, minority or refugee groups, but also violent non-state actors such as armed groups, rebels and terrorists (Mihr 2012). Among non-state actors we also find academics, experts and former policymakers and legislators who speak only in their private capacity and do not do so on behalf of or to represent any institution, whether it be a state or a non-state institution.

In our three-year, interdisciplinary research project, ‘The Impact of Transitional Justice in Democratic Institution Building’, which covered eight countries throughout the world, the challenges were to coordinate field trips, interviewers, language interpreters and both state and non-state interviewees at different stages and phases of the project. Some interviews required interpreters while others did not. Some were conducted partly with interpreters during the first months of our research in Sierra Leone and Brazil. Other interviews were conducted only after 20 months into the research project in countries such as Germany, Japan and Hungary. In our project, a structured questionnaire would have made collecting the data we needed very difficult because, over time, new questions emerged, earlier ones were dropped and new ones were included in the interview guide. A number of researchers made several visits to the countries in question and conducted semi-structured interviews. Strict questionnaires would have generated

only ‘safe’ facts and figures that had already appeared elsewhere in the literature. Semi-structured interviews with an interview guide and clear ‘why’ questions proved to be the best approach to collecting a rich set of qualitative data to be assessed and to provide answers to our main research question. A guide, such as the one we used, adapts to the dynamic and changing nature of a project and its focus over a period of time.

Thus, this chapter focuses on examining the use of semi-structured interviews as viewed through different empirical research projects that involved non-state security actors, including CSOs, NGOs and experts from business, law and academia or human and civil rights activists or separatist-group sympathizers who have responded in their private capacity but have been part of a regime transformation or of decision-making processes concerning a political event (e.g. the making or implementation of laws).

The purpose of semi-structured interviews with non-state security actors

In a research project, semi-structured interviews are suitable for addressing ‘why’ questions and generating qualitative data that allow for further interpretation. In the social sciences and the humanities, they are the core of empirical research designs of this type (Bernard 2006). Interviewing key stakeholders adds to the qualitative component of the design by unravelling and uncovering information, data and nuances of policymaking of both the recent past and the present. By juxtaposing quantitative data with the views and positions of stakeholders, as seen from the perspective of the interviewees and respondents, this approach contributes to a triangular research design. Non-state actors and experts are of particular interest because they often talk about their personal experiences and motivations, as well as about the failures in their work – that is to say, unlike state officials or authorities, who represent institutions and are less likely to share their own experiences, non-state actors often respond in their personal capacity as independent individuals (Whiting 2008; Mihr 2011).

In human rights, post-conflict and transformation research, non-state actors are essential respondents, because they add value to the institutional knowledge and data that are already available. In post-conflict societies and countries in transition, non-state security actors in the field are often key to establishing or re-establishing security. For research projects conducted during this period (i.e. the period of transitional justice or institution building), security actors, such as former combatants or members of armed groups, child soldiers, and paramilitaries, or agencies that provide security training and capacity building (i.e. in the context of demilitarization) are of central interest to the researchers. They tell the stories behind the ‘big’ political agreements and decisions that are often analysed through quantitative research.

In contrast to the non-state armed security actors, active members of grass-roots organizations or human rights NGOs, such as Amnesty International, will relate personal experiences such as their successful (or unsuccessful) lobbying

efforts or will describe how they finally managed to make themselves heard and actually change the political agenda by participating in parliamentary hearings on torture or on refugees. In political science research, NGOs, CSOs, private enterprises and businesses, as well as individual activists and foundations, have often turned entire political agendas around, which is why they deserve our attention when we are conducting empirical and qualitative field research.

The best way to approach these groups is through semi-structured interviews, which leave room for personal reflections, storytelling and digressions, because they are guided by a major ‘why’ or ‘how did it happen’ question. The interviewer’s role is to guide the respondents through the talk in order to help them remember and share information that will supplement quantitative data. For example, we can quantitatively assess how many women are members of parliament in a particular country, and we may also know that this number was once regulated based on a political decision allowing a specific quota. But in order to know who pushed and lobbied for this quota, we need to conduct semi-structured interviews, particularly with non-state actors (such as female members of NGOs) who might have been the driving force behind such efforts.

Quantitative data can thus be supplemented by qualitative data based on observations, participation and face-to-face or online interviews. The purpose of these types of interviews is to provide further insights into motivation or behaviour (e.g. asking NGO representatives why they made or did not make specific decisions), into the political dynamics, systems and decision-making processes in non-state armed or security groups, NGOs, CSOs, and into the relationships among actors and among institutions. One might ask, for example, whether citizens do or do not participate in elections and whether they trust parliaments and governments or courts (Horton *et al.* 2004). The main criterion used in choosing this interview method is to understand ‘how’ and ‘why’ organizations and institutions do what they do and their underlying motivations; in other words, it generates pieces of information that would not necessarily be available through observations or written statements from these groups. Thus, the data and results generated by these interviews must be considered complementary to other available data.

In social science research, semi-structured interviews are most valuable when the subject matter has a strong attitudinal and behavioural focus. This is the case, for example, if one wants to determine why some institutions receive more trust from citizens than others, why some democracies function better than others and why people participate in grass-roots organizations more actively in some societies than in others (Trinczek 2009). Because these institutions are legitimized through citizen participation, it is important to interview those groups – namely, citizens and non-state actors – who ‘decide’ whether to trust an institution or not, whether to participate in elections or civil rights movements or not, and so forth.

In International Relations or comparative politics, the aim of semi-structured interviews is often ‘to look behind the scenes’ in order to gain insights into the motivations and factors that led to the adoption of certain international human rights treaties, such as the Charter of Fundamental Rights of the European Union

of 2000 and the UN Convention on the Rights of Persons with Disabilities of 2006. In this case, such interviews could be used to obtain information about the procedures and dynamics that led to a certain outcome (e.g. a human rights treaty) as a way of filling gaps in previously documented data, such as one would find in academic articles, books or databases (Horton *et al.* 2004: 339).

During transition periods, non-state security forces, such as paramilitaries and private security companies, often fill the security gap left by the government's inability to provide protection. However, these forces often share different reasons for laying down their arms, demobilizing and becoming civil security services for the time of transition to a functional governmental security sector. For researchers in the area of post-conflict or transitional justice, the reasons security actors relinquish their arms and become civil servants are of primary interest, not least because knowledge of these reasons allows researchers to draw direct links to how institutions are built during such periods.

Another purpose of semi-structured interviews is to verify or falsify critical assumptions and hypotheses concerning larger groups, peers or target groups, some examples being children, women, minorities or persons aged 60 years or older. Surveys are one way of finding out how the population at large or specific target groups think and act (see Chapter 13 by Sabrina Karim). Semi-structured interviews can also help find out about people's motivations; they allow us to explore the extent to which our theoretical priors are reflected in the actual behaviour and perception of significant actors and enable us to formulate new grounded theories (Horton *et al.* 2004: 348).

When the first survey by Sommer and Stellmacher (2009) on human rights awareness in Germany was published in 2001, the results were surprising because the percentage of citizens who had a relatively broad understanding of human rights was unexpectedly low, and the few who did know about human rights said they had learned about them through their interaction with Amnesty International. But to determine why and how this was so, and why an NGO and not the state education system was the source of this human rights knowledge in Germany, it was necessary to interact with these people by means of semi-structured interviews. The answers were revealing in that it became clear that the formal education system in Germany had failed to provide systematic education about human rights because those in charge were under the misapprehension that political education about state institutions is the same as human rights education (Mahler *et al.* 2009). Semi-structured interviews are also valuable in creating a network of contact persons for the researchers on individual levels that can later be a source of (follow-up) information, unlike structured interviews, which do not allow for such personal networks and ties to emerge from the interviews.

Another category of non-state interviewees are those we often refer to as 'experts'. In such cases, it is important for the interviewer to adapt to the background of the expert, whether that person has a managerial, humanitarian or military background. For example, private security officers, managers, academics, human rights consultants or former military officials, even when they are outside their office settings, tend to conform to the habits of their working

practices and rhetoric. For example, when respondents are asked about their experiences in their (former or current) capacity as CEO of a company or as a former police officer or civil rights activist during, say, the political transition in Hungary or Brazil in the 1990s, they will automatically revert to behaviour typical of their working situation at that time.

The success and added value of semi-structured interviews with experts therefore depend largely on the social competence of the interviewer, not on that of the respondent. Many of these experts are oriented to predominantly question-and-answer interactions. Interviewers run the risk of dooming the whole interview to failure if they begin by inviting a respondent to engage in a lengthy narration, because the semi-structured interview setting is often diametrically opposed to the everyday communication structures that characterize the expert's workplace (Trinczek 2009: 205–206). Thus, changing the nature of an interview environment to one that would be more appropriate for a dialogue format can require considerable time and social competence on the part of the interviewer.

In relative terms, talking to non-state or state experts during the exploratory phase of a research project is a more efficient and focused method of gathering data than is relying on participatory observation or systematic quantitative surveys. Conducting expert interviews can shorten time-consuming data-gathering processes, particularly if the experts are seen as 'crystallization points' for practical insider knowledge and are interviewed as surrogates for a wider circle of players. Expert interviews are also called for when one's access to data might otherwise prove difficult or impossible (Bogner *et al.* 2009: 2), which is particularly true of non-state experts who are not bound by a particular state doctrine or by an official narrative of past events.

Generally speaking, due to the growth of multi-stakeholder participation in major events, both private and public and on the international level as well as on national levels, the role of non-state actors becomes increasingly important for political action and thus for political science research. In this context, the variety of actors and the question of whom to consider either a state or a non-state actor become increasingly blurry. This question and the choice of one interviewee over another should be carefully investigated and assessed before an interview is conducted.

Methodological challenges

There is always the risk that qualitative research and interview data may be over-interpreted. Although semi-structured interviews allow us to interpret people's behaviour, opinions and motivations, they can also turn into storytelling sessions, which provide little in the way of analysable data. Often, interviewees/respondents will invite the interviewer to accompany them to a private meeting, an event or a political caucus or manifestation to witness (and thereby verify) their estimations of a situation first hand. This option must be cautiously assessed by the interviewer, because too close a relationship between the two parties can result in bias when the collected data are being interpreted. Although such

participatory observations may be useful in interpreting the respondent's spoken word and other data, accompanying and spending additional time with the interviewee and other members of that particular peer group in a more social environment might in some ways influence or manipulate the interviewer in terms of the interviewee's/group's motions, moods, personal affinities or aversions. Depending on the circumstances and the non-state actor, it may be the interviewee who, consciously or not, selects the location of the interview. S/he can dominate the interview and manipulate, for example, which question might come next or whether to let the flow of the narrative continue as it is (which would not happen in a structured interview). On the other hand, by (and when) guiding the respondent through the interview, the interviewer influences the process and outcome of the interview in a significant manner.

Selection of actors and sampling of groups

Another challenge is choosing interviewees on the basis of specific criteria, such as level of education (Barriball and While 1994: 328–329). Non-state and security actors are by no means a homogeneous group, ranging from the illiterate to the highly academic. When it comes to state actors, however, it is highly unlikely that someone at the ministry of defence or a police academy would not be fairly well educated, whereas in the realm of non-state actors, the chances are much greater that one will be asked to interview people with lower levels of education. Of course, a person's level of education does not necessarily reflect their intellectual capacity, but it does have an influence on their ability to be articulate during an interview. This factor can, in turn, influence the way we ask questions and guide or manipulate respondents throughout an interview. When I conducted interviews with members of the Roma and Gitano communities in Slovakia and Spain, respectively, about whether they felt their human rights were respected or not, I met many people who had a low level of education but were eager to be interviewed (Mahler *et al.* 2009). They certainly had opinions, but their language and their ability to express themselves had to be taken into account and put into context. In such cases, the interviewer must keep in mind a certain responsibility to the interviewee. The same is true when conducting face-to-face interviews with children. It is often tempting for an interviewer to abuse the situation and twist interviewees' words in a way that will make their responses consistent with the interviewer's view but not necessarily with the interviewee's real views.

Non-state actors vary widely, ranging from eyewitnesses to peer groups, from representatives of NGOs to grass-roots activists to armed security actors, such as rebels. They can be academics, experts or former police or military officers. This broad category also includes actors who are difficult to contact, such as former guerrillas, warlord combatants, terrorists, neo-Nazis and members of other radical and violent groups and organizations. But, in any of these cases, the success of attempts to persuade these actors to agree to be interviewed will depend on the interviewer's social competence and his or her ability to seek

these actors out. Even though the interviews are conducted for scientific purposes, actors of these backgrounds often will agree to interviews only under the condition that they will have the opportunity to send out their political message. Researchers must be aware of the political and even radical messages that an interviewed actor might want to send out by using the interview for his or her own benefit and to justify his or her violent actions. At the latest when we need written approval and authorization of the interview for scientific ethical purposes (usually by a written consent form), we will see that many actors withdraw their consent for us to cite passages from the interview in our publications.

Some years ago, for my research on transitional justice in Spain, I interviewed supporters of terrorists affiliated with Basque Homeland and Liberty (ETA). These supporters agreed to be interviewed only on the condition that their names and personal affiliations would be anonymized. For this reason, the material in these interviews (or, in this case, talks) could not be used for interpretation, but it did serve as valid background information that could be assessed and it provided some context for the other interviews and data in this research project (see Mihr *et al.* 2011). It goes without saying that it is in the respondents' own interest to provide their version of events to the interviewer in the hope that it will be cited in an academic publication. However, interviewing radical actors is a sensitive issue; by agreeing to be interviewed, they often use the platform offered by the academic research unit – namely, the interviewers – to justify their violent actions and quest for legitimacy (Clapham 2006).

A similar example can be found in security or transition research when one is conducting interviews with victims and survivors of imprisonment or mass atrocities such as genocide. During my research on human rights in Germany and later in Ruanda, I met not only many political prisoners, victims of torture and survivors of genocide but also some of the perpetrators (Mihr 2013). Asking intrusive questions in such cases can be traumatizing to the respondents and lead to unpredictable responses, such as weeping or complete silence. At the very least, researchers in these difficult positions should prepare for such eventualities by talking to psychologists who know how to approach traumatized survivors. I have met interviewers who could not forgive themselves for subjecting an interviewee to such a stressful situation. Apologies afterwards are often not enough. Even though many victims and survivors are now organized in support groups and can be contacted through NGOs that offer internet access or visibility (which is why we find them), they are not necessarily strong enough psychologically to face an academic researcher who asks straightforward questions concerning details of past atrocities, which the interviewees may have forgotten or mentally blocked until they are reminded as we seek answers to the why and how of their violent past.

Interviewers must carefully select the activists and members of NGOs or CSOs to be interviewed, especially those on the local or grass-roots levels. For example, in researching the impact of transitional justice in Uganda, a country whose administrative language is English, I interviewed victims and survivors who had endured a decade-long war and atrocities committed by the Lord's

Resistance Army in the northern part of the country. These interviews were challenging in a number of ways. One first had to overcome the language barrier. In addition, the women refused to be interviewed alone, and they were not allowed to speak without a man being present; tribal or village leaders (usually men) would speak on their behalf or try to find someone in the village who could speak English, sometimes a young man with some high-school education, even though he himself had not experienced the war. These leaders or spokesmen often spoke on behalf of the entire village. The fact that the level of education, traditions, gender and other ethical or religious issues impeded face-to-face interviews was not the most troubling aspect – in general, the problem is rather the fact that, as researchers, we might not be self-critical enough about such interviews. We are easily tempted to improperly consider responses to be legitimate sources of data as a way of verifying hypotheses that otherwise could not be corroborated. In short, the problem does not lie with the interviewee because of linguistic, educational, religious or gender differences or because of the peer group as such; rather, the problem lies in how we generate and reflect on the information and data they share with us.

Regardless of these difficult circumstances, there is one advantage to conducting face-to-face interviews; that is, the opportunity to validate the respondent's answers by observing non-verbal cues, particularly when sensitive issues are being broached. We often refer to this type of attention to non-verbal indicators as 'participatory observation'. When interviewing pairs or groups of people, we must ensure that all questions are answered by all interviewees, but we should also take note of the respondents' behaviour and the surroundings while trying to prevent the effects of group pressure on their answers (Nasiritousi *et al.* 2016). Interviewing members of these groups in sequence, if possible, is usually better, because it makes it possible to compare the answers later.

However, most of the positive responses to our requests for an interview come from persons with academic degrees or from academics and researchers themselves. Not only do they sympathize with us as researchers, but they also feel less threatened, have fewer language issues and, last but not least, are happy to 'advertise' their latest publications, articles or policy briefs throughout the interview as a way of promoting their own academic or activist work in the area at hand.

The second group that often agrees to be interviewed includes leading representatives of NGOs or CSOs that depend on international or public donations and funds – the 'donor-driven' NGOs. They usually speak a common world language (mostly English) because they must regularly report to their external donors, most of which are located in the Global North – namely, in Europe or North America. That is why many of these representatives are fairly proficient in English and agree to participate. Because they are under constant pressure to report and be accountable to their (international) stakeholders for the work they do (and the money they spend), they are more available for interviews. Being quoted and cited in researchers' publications or on their websites is a form of reference and accountability that may attract future donors. I mention this here

because, although we all acknowledge that many representatives of NGOs work hard and earn little, we should all be realistic enough to know that they do not consider it a priority to spend time with researchers from universities based mainly in the Global North unless they can expect some sort of benefit for themselves.

A third group of non-state actors that tends to respond quickly and positively to interview requests includes former officials, leaders or representatives of NGOs, political parties or state authorities who are now willing to participate in their capacity as private citizens and to whom we often refer as experts or eye-witnesses. In this context, language and memory difficulties are some of the challenges, which occurred, for example, when interviewing civil rights activists or former members of the *Solidarnosc* movement that was active in Poland in the 1980s. These activists sometimes speak only in their mother tongue (in this case, Polish), and they may combine their own memories and experiences with narratives heard from participants or the media after the fact and which do not necessarily reflect their own personal experience.

Before I proceed to discuss some other methodological challenges one may face during semi-structured interviews with non-state interviewees, I would like to share one more observation. When doing research in our home countries, we rarely confront language problems unless we are interviewing linguistic minorities or refugees, and, when going abroad, we often choose countries in which we are able to communicate in at least one of the common administrative languages, such as Spanish in Latin America, French in West and North Africa or English in East and South Africa and India, and so on. Thus, we may unintentionally exclude emerging countries such as Belarus, China, Iran and Kazakhstan in anticipation of potential language difficulties. At the same time, a lot of the grass-roots organizations and non-state actors in such countries depend on international financial support (as explained above). As a result, they need to translate the native-language websites into English, French, Spanish or Arabic, depending on their major donor country/countries, thus facilitating contact prior to our field visit. By contrast, NGOs that do *not* translate their websites into at least one of the more common global languages will rarely be ‘detected’ by our research team (unless we have grant money to pay a native speaker or colleague who knows the language of a specific country). We should be aware of these drawbacks and recognize that language barriers can be an obstacle and a source of bias in our research, as can be non-state actors supported and driven by external donors (Fontana and Frey 2005).

Similar problems may arise when we conduct interviews with non-state actors in wealthier countries, such as South Korea, Germany and Japan in which none of the world’s languages is a common one. Because such actors do not depend on international donors, instead receiving all their support from within their own countries, it can be almost impossible to find civil rights activists, victims or eye-witnesses who speak English, even though they are well organized and do not suffer from a shortage of money. I know this from personal experience, having tried to arrange for interviews about transitional justice with non-state actors or victims of a dictatorial regime in South Korea and elsewhere. The NGO or CSO

websites appear only in the language spoken by the particular nation because they are accountable only to their own national donors or agencies; they have no need to promote their cause in English because it would offer no political and monetary benefit. In such cases, we had to hire native speakers to help us examine Japanese, German or Korean websites to identify potential interviewees. Among those interviewees were civil rights activists who had been leaders in the student movement during the June Democratic Uprising in South Korea in 1987, victims of Stasi repression under the former Communist regime in East Germany, and NGO activists in Japan who initiated reconciliation programmes with former slave labourers and inaugurated private memorials for victims of the Second World War throughout the country. Had we not invested in translation services, we would have missed the opportunity to meet with these interviewees and obtain a considerable amount of valuable data.

The same is true when conducting research in European or Asian countries where English is not a common language, such as Poland, South Korea, Japan and Hungary. In countries such as these, most non-state actors publish websites in their respective language. This limitation impedes our preparatory research and biases the research design if we are unable to select equal pairs and sets of non-state actors.

The essential link of the ‘why’ question

The semi-structured interview is an asset that allows us to make a link between what people say they did and what they actually did. The data collected capture the nuances and sensitivities of people’s actions and the meaning of such actions in different contexts – something that purely quantitative data does not permit. Data generated through interviews help to capture how people perceive their actions or their relationship to institutions, regime change, neighbourhood policies, and so forth. Such personal knowledge is indispensable to a thorough assessment when one, for example, investigates regime change or creation, conflicts or peace. To understand why, in 1992, the European Economic Community decided to establish a political union and its member states formed the European Union, we have to interview contemporary stakeholders (now experts) and eye-witnesses of that time. Publicly available ‘facts’ and quantitative data are not necessarily profound or accurate enough to explain the momentum, the motivating factors, and why, after the end of the Cold War in 1990, these countries were ripe for a political union. Without qualitative knowledge and data through interviews, it is difficult to answer the big ‘why’ questions: for example, why did the institutions, governments and citizens of the European Union act or react in certain ways? Thus, through semi-structured interviews, we can better understand major political and societal changes by examining their micro-foundations and the reasons they take place, with the added value of being able to focus on their attitudinal and behavioural aspects.

The method of semi-structured interviews helps explain multi-causal or cumulative causal relationships. This holds regardless of whether the research

design is multi- or monodisciplinary. Interviews oriented toward qualitative research may seek to obtain both factual information and the meaning of such information. However, conducting an interview to extract meaning or purpose is usually more difficult owing to the subjectivity of the interviewer (and the interviewee), as explained above (Kvale 1996). The interviewers' main responsibility (and the main challenge they face) is therefore to control their own subjectivity and be as neutral and reflective as possible in trying to understand the meaning behind the interviewees' words in the most unbiased and objective way possible (Horton *et al.* 2004: 340).

Edwin van Teijlingen (2014) divides semi-structured interviews into four phases: (1) the planning phase; (2) the doing phase; (3) the analysis phase; and (4) the reflecting phase. These four phases will now be described in more detail.

The first phase, planning, is when the questions and the 'interview guide' containing the 'why' question are formulated. It should also make clear what motivated the researcher to ask these questions and thus refer to the general design and goals of the research. During this phase, one should already be preparing the interview and establishing a communicative atmosphere.

The second phase, doing, consists of the actual interviewing in the field with the interviewee. An interview can last from 30 minutes to several hours. The researcher/interviewer has to listen carefully to the respondent and observe any non-verbal behaviour as well. An interview can be very different depending on where it takes place (e.g. in an office space, a public cafeteria or restaurant or a private home). Non-state actors often prefer either a public space or a private home for the simple reason that their office is too small. My best interviews are usually conducted in private homes, where there is no external interference and the interviewees soon tend to warm up, respond more freely and engage in narratives that would otherwise never have come to light. Even though the respondent may not be willing to authorize certain stories later, additional data may prove useful in putting these stories and emotions into context; conversely, these stories may help us interpret the data in a new light. In addition to the interviewees' actual, authorized answers, their verbal and non-verbal behaviour can be captured during this phase and be useful in evaluating their responses.

The third phase, analysis, involves creating, transcribing and/or summarizing the interview records. It is also the phase during which the interviewers should analyse their own behaviour to determine whether answers that otherwise would not have been given were somehow triggered or to review whether they had risked retraumatizing interviewees who had been victims or were survivors of horrific acts, as explained above. Researchers often only realize afterwards how strongly they may have manipulated or influenced the outcome of the interview as a result of uncontrolled reactions, suggestive or polemic questions, mimicking or persuasive body language.

Therefore, during the fourth phase, reflection, the researcher should identify information gaps in the interview and also prepare for the next interview. This phase allows us to add new questions generated by the previous interview(s) or subsequent discussions that can be addressed in the forthcoming interview(s)

with the same or a different person. Researchers who work in an interdisciplinary environment for a long period of time and meet a variety of interviewees and interviewers may have more opportunities to learn from previous interviews. Individual researchers, such as PhD candidates and postdoctoral fellows, may wish to return to interview the same candidate to obtain information they overlooked during the first interview; however, non-state actors often have limited resources and time and may not necessarily agree to be interviewed by the same researcher more than once.

In general, interviewers who plan to conduct a semi-structured interview are advised to listen more than to speak, because most interviewers tend to talk too much. It is important to ask questions in a straightforward, clear and non-threatening way in order to avoid distracting the interviewee, and it is crucial to avoid using phrases such as ‘I think that...’ or expressing personal opinions through cues and hints, because interviewees may then try to please the interviewer by giving what they believe are correct or desired responses. Finally, the interviewer must show genuine interest in what the interviewee is saying and appear to be enjoying the exchange; we should conceal any feelings of fear or boredom by maintaining a consistent tone of voice, posture and facial expression (van Teijlingen 2014).

Thus, prior to an interview, interviewers should have acquired in-depth knowledge of the data, facts and figures about the subject to be addressed. The interviewer must guide the respondent through the interview and be as neutral as possible, drawing on extensive knowledge of the case under investigation. For example, in my interviews with eyewitnesses and civil rights activists involved in the political transitions in East Germany and Hungary in the 1990s, I discovered that the interviewees were often unaware of facts about the period in question and confused their own experience with stories they had heard or read. For this reason, researchers must always carefully confirm and filter interview content by checking it against existing external data and established information. In particular, we should examine how interview data are shaped by general historical narratives and popular belief, the social environment and the research process itself.

Probing too much, too early or too late is another mistake one can make when conducting semi-structured interviews, particularly with non-state actors and their representatives who are not used to giving high-level academic interviews or have language impairments. Although well meaning, we may interrupt an interviewee who is getting carried away with a story, memories and narratives or who cannot find the right words to express emotions and memories. In trying to ‘help’ in this way, we interrupt the interviewee’s train of thought, possibly missing out on important information. To listen and to be patient are probably the most important skills in this type of interview (Fontana and Frey 2005). Probing can be verbal or non-verbal and, to be sure, can help direct an interview and ease the respondent back on track. To accomplish this, one can pose questions such as ‘Can you give examples...?’ and ‘How do you feel about this...?’, or one can simply ask for justification of what has been said with a question such

as ‘Could you please tell me in your own view the pros and cons of this particular situation and why you think that is the case...?’ Trained interviewers listen attentively and only prompt an interviewee when it is absolutely necessary.

Because most of the non-state security actors we interview are eyewitnesses, activists, representatives or experts, they are agents of different rationality models and have been effective in that they actually succeeded in influencing public awareness through the media or politics. Political scientists seek to interview these individuals because they are an integral part of the subject matter we wish to explore, whether it be in International Relations, comparative politics, regime assessments or contemporary history. Bogner *et al.* (2009) argue that the ability to put specific knowledge to use for political gain is a constitutive characteristic of this type of ‘post-traditional’ expert (Bogner *et al.* 2009: 4). For an expert or a counter-expert, therefore, the goal is to interpret the world in a high-profile and influential but not necessarily new manner and thus become a powerful voice in the struggle for the establishment of certain conditions of definitions. From a methodological perspective, expert knowledge is crucial, so interviewers must be aware that experts are on a constant ‘mission’ to promote and articulate their views to a larger audience and not necessarily to reveal a truth or to help the research project to be a success.

Thus, key methodological challenges for semi-structured interviews with non-state actors are the selection and identification of actors and their role in the research project (Mihr 2011). Gender, language or cultural differences play a greater role when one interviews non-state actors as opposed to state actors, because state actors give interviews based on their position or have received clearance from their superiors to participate in their official but not private capacities. Moreover, they tend to have interpreters who can facilitate the interview. We interview those actors because of their position within the institutional framework they represent, not so much because of their personal motivations, experiences and future visions, which can be very personal and private. Non-state actors are more sensitive to these questions, because they do not represent a formal or state agency and often act outside the formal or legal governmental framework.

Preparing well for the interview and obtaining as much information and factual data about the situation, era, incident or event of interest as possible are of the utmost importance in being able to separate the respondent’s story from external data. Interviewers should avoid being too well acquainted with the interviewee but at the same time should create an atmosphere that is comfortable for both parties. The interview guide containing the ‘why’ question should constantly be adapted as the research progresses. The different phases – preparation, interviewing, analysis and reflection – distinguish the semi-structured interview from the structured interview, which involves only a question-and-answer method. The success of a semi-structured interview will be evident when the researcher feels that she or he has been able to obtain information and data that explain and answer the ‘why’ and ‘how’ questions that are otherwise not available in the existing literature or databases.

Conclusion

The take-home message for those who conduct field research that involves semi-structured interviews with non-state actors is that it is a learning process and a skill that anyone with open research questions can acquire. Regardless of whether the research project is comparative, multicausal or monocausal, what makes interviewing non-state actors an asset is the fact that we generate data that help explain the circumstances and the reasons for the subject matter under study. Generally speaking, non-state security actors are not bound by narratives based on official state doctrine when answering our questions. During the interview, they are more flexible, and they are willing to be guided and to accept in-depth questions and discussions.

Contemporary empirical field research and interviews are focusing more and more on non-state actors and less on state actors for two reasons. First, when interviewing state actors, such as representatives of governments or diplomats, the outcome has often been disappointing because the interviewee later failed to authorize the responses if they did not follow strict state orders or because the answers were ones that could have been found on the website of the respective institution, such as parliament or the ministry of justice. Non-state and private corporate actors, particularly in the emerging realm of multi-stakeholder decision-making (e.g. in cyber security and governance), are being asked to assume state duties and responsibilities to safeguard the security of citizens or internet users. This is an area that is gaining significance for political science researchers (Nasiritousi *et al.* 2016). However, the methodological challenges remain the same regardless of which actor is interviewed, because the success of an interview depends mainly on how well the interviewer is prepared and whether the interview is conducted in the most neutral and professional way possible.

References

- Barriball, K. Louise, and Alison While. 1994. 'Collecting Data Using a Semi-Structured Interview: A Discussion Paper'. *Journal of Advanced Nursing* 19(2): 328–335.
- Bernard, H. Russel. 2006. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. 4th ed. Lanham: AltaMira Press.
- Bogner, Alexander, Beate Littig, and Wolfgang Menz. 2009. 'Introduction: Expert Interviews – An Introduction to a New Methodological Debate', pp. 1–13 in *Interviewing Experts*, A. Bogner, B. Littig, and W. Menz, eds. Houndmills: Palgrave Macmillan.
- Clapham, Andrew. 2006. *Human Rights Obligations of Non-State Actors*. Oxford: Oxford University Press.
- DiCicco-Bloom, Barbara, and Benjamin F. Crabtree. 2006. 'The Qualitative Research Interview'. *Medical Education* 40(4): 314–321.
- Fontana, Andrea, and James H. Frey. 2005. 'The Interview: From Neutral Stance to Political Involvement'. pp. 695–728 in *The SAGE Handbook of Qualitative Research*, 3rd edn. N.K. Denzin and Y. Lincoln, eds.. Thousand Oaks: Sage.
- Horton, Joanne, Richard Macve, and Geert Struyven. 2004. 'Qualitative Research: Experience in Using Semi-Structured Interviews', pp. 339–357 in *The Real Life Guide*

- to *Accounting Research: A Behind-the-Scene View of Using Qualitative Research Methods*, C. Humphrey and B. Lee, eds. Amsterdam: Elsevier.
- Kvale, Steinar. 1996. *Interviews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks: Sage.
- Mahler, Claudia, Anja Mihr, and Reetta Toivanen, eds. 2009. *The United Nations Decade for Human Rights Education and the Inclusion of National Minorities*. Frankfurt am Main: Peter Lang.
- Mihr, Anja. 2011. 'Non-Governmental Organizations in the Human Rights World', pp. 377–390 in *The Ashgate Research Companion to Non-State Actors*, B. Reinalda, ed. Farnham: Ashgate.
- Mihr, Anja. 2012. 'Non-State Actors in Conflict', pp. 305–323 in *Human Rights and Conflict: Essays in Honour of Bas de Gaay Fortman*, I. Boerefijn, L. Henderson, R. Janse, and R. Weaver, eds. Antwerp: Intersentia.
- Mihr, Anja. 2013. 'Transitional Justice and the Quality of Democracy'. *International Journal of Conflict and Violence* 7(2): 298–313.
- Mihr, Anja, Jos Philips, and Isabel Duyvesteyn. 2011. 'Human Rights NGOs: Imperative or Impermissible Actors in (Post)-Conflict Societies?' *International NGO Journal* 6(11): 229–239.
- Nasiritousi, Naghmeh, Mattias Hjerpe, and Bjoern-Ola Linnér. 2016. 'The Roles of Non-State Actors in Climate Change Governance: Understanding Agency through Governance Profiles'. *International Environmental Agreements: Politics, Law and Economics* 16(1): 109–126.
- Polit, Denise F., and Cheryl T. Beck. 2006. *Essentials of Nursing Research: Methods, Appraisal, and Utilization*. 6th edn. Philadelphia: Lippincott Williams & Wilkins.
- Sommer, Gert, and Jost Stellmacher. 2009. *Menschenrechte und Menschenrechtsbildung: Eine psychologische Bestandsaufnahme*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Trinczek, Rainer. 2009. 'How to Interview Managers? Methodical and Methodological Aspects of Expert Interviews as a Qualitative Method in Empirical Social Research', pp. 203–218 in *Interviewing Experts*, A. Bogner, B. Littig, and W. Menz, eds. Houndmills: Palgrave Macmillan.
- Van Teijlingen, Edwin. 2014. *Semi-Structured Interviews*. Bournemouth: Bournemouth University Graduate School. Available from: <https://intranetsp.bournemouth.ac.uk/documentsrep/PGR%20Workshop%20-%20Interviews%20Dec%202014.pdf> (accessed 24 October 2016).
- Whiting, Lisa S. 2008. 'Semi-Structured Interviews: Guidance for Novice Researchers'. *Nursing Standard* 22(23): 35–40.

6 Combining semi-structured interviews and document analysis in a study of private security expertise

Joakim Berndtsson

Introduction

This chapter introduces the reader to a study of security contracting by the Swedish Ministry for Foreign Affairs (MFA). The study was conducted as part of a larger research project entitled ‘Security for Sale’ which I worked on together with my colleague Professor Maria Stern between 2009 and 2013. The overarching aim of the project was to map out and analyse different understandings of security work and expertise among public and private security professionals. Originally, the idea was to research security privatization in several different countries. As the project progressed, however, we grew increasingly interested in the Swedish case, having noticed significant changes in terms of the official view of – and indeed use of – private security among state agencies. In the end, the project resulted in several case studies within Sweden, covering a variety of aspects of security privatization, such as the use of private security companies (PSCs) at Stockholm Arlanda Airport, the implications of security privatization for Swedish security and development work in Afghanistan, the view of PSCs among Swedish military officers, and, as will be explored below, the contracting of PSCs by the Swedish MFA for diplomatic protection (Berndtsson and Stern 2011, 2013; Berndtsson 2012, 2013). With the exception of the study on Swedish ‘PSC-military relations’ (Berndtsson 2013), the studies relied chiefly on interpretative or qualitative analyses of textual sources collected through interviews and official state documents.

The study discussed in this chapter was initiated in 2010 and subsequently published in *Millennium: Journal of International Studies* as ‘Security Professionals for Hire: Exploring the Many Faces of Private Security Expertise’ (Berndtsson 2012). The study was prompted by the decision by the Swedish MFA in 2008 to contract with private security contractors to provide armed protection and security management services at the Swedish embassies in Iraq, Afghanistan and Sudan. Before this time, the official Swedish position on the use of PSCs in high risk or conflict areas had been characterized by reluctance and suspicion, but this changed into a much more active engagement with, and a growing reliance on, commercial security services. From our perspective in the research project, this change raised a number of pertinent questions: Why had the

Swedish position on private security changed? How did PSCs become seen (and rationalized) as accepted and trusted security providers? What were the consequences of this development for Swedish engagement in peace operations and in development assistance and reconstruction efforts? While all of the questions were deemed interesting and relevant for the project, this particular study needed a more narrowly defined focus. With the overarching aim of the research project in mind, it was decided that this study should focus primarily on one particular question, namely: How were images of private security contractors as 'security experts' articulated and received by the key actors in the MFA contracting process?

At this point in time, fairly little was known about the Swedish use of PSCs beyond the basic fact that both the Swedish Armed Forces and the Swedish Ministry for Foreign Affairs had started contracting with private security companies as early as 2005/2006.¹ However, there was a general lack of knowledge about the nature and scope of contracts, the use of public procurement procedures or the views of contracts and companies among state security officials. This meant that the study needed to map out the basic facts of the case as well as to collect data that would make it possible to address the main research question. In the end, a qualitative, exploratory study was designed, focusing empirically on the MFA contract and relying on a combination of data collected from official state documents as well as semi-structured (elite) interviews with both PSC and MFA representatives. The method for analysis was 'thematic' in the sense that data were systematically coded and scanned for key themes and concepts and to examine how they were used in the construction of private security expertise.

In this chapter, the aim is to describe and reflect on the ways in which this study was designed, how and why the different methods were chosen and used, as well as the methodological implications of these choices. Importantly, it will also discuss a number of practical problems that arose in relation to the organization and carrying out of the study, and the steps taken to overcome them. The chapter proceeds as follows: the following section will describe some of the key theoretical points of departure for the study and how these, together with the research question, provided a basis for the design of the study and the choice of methods for data collection and analysis. In the following section, the application of methods will be discussed in more detail, focusing chiefly on strategies for, and problems with, data collection. Finally, the concluding section will reflect on the experiences of applying these methods for studying private security actors more broadly.

Designing a study of private security professionals

As explained above, the aim of the study of the Swedish MFA contract was to find out how PSCs are constructed and accepted as experts on security. This issue is important because it links up with larger questions about the ways in which state and non-state security actors and activities are legitimized in different contexts. Theoretically, the study draws on the (largely constructivist and

critical) work on security expertise, risk management and security knowledge by scholars such as Anna Leander and Elke Krahmann, as well as work on 'security assemblages' by Rita Abrahamsen and Michael C. Williams and on private security identities or corporate cultures by Kateri Carmola (Leander 2006, 2007; Krahmann 2011; Abrahamsen and Williams 2011; Carmola 2010). A key argument is that PSCs are 'hybrid' actors who draw on multiple 'identities' or 'cultures' to construct images of security expertise and professionalism. This construction or messaging is of course done in relation to (and communication with) several different audiences, such as prospective clients, or other security actors and professionals. In this context, Carmola has helpfully pointed out that PSCs frequently draw on, and shift among, different organizational cultures or characteristics depending on the context and the audience, including the 'worlds of the military, the business world and the humanitarian NGO' (Carmola 2010: 28).

From this perspective, it becomes clear that in order to get at the formation of private security expertise, and to capture the hybrid and shifting nature of PSC 'identities', it is necessary to study not only the companies' articulation of self-images in different settings and the ways in which they draw on various forms of specialized and legitimate knowledge, but also to study the ways in which these are received by different audiences. Doing so might generate interesting new knowledge about the complex and variable nature of PSCs, thus moving away from oversimplified, static and potentially misleading descriptions of what PSCs are, and how they become legitimate non-state security actors. In line with this argument, the study of the Swedish MFA contract focused both on how Swedish state authorities describe and specify the kind of security knowledge and skills they want from a private security contractor, as well as the ways in which the PSCs involved in the contracting process present themselves as security professionals or with expert knowledge in different contexts, for instance as they strived to meet the demands of (and thus become accepted by) the prospective client.

The study was thus designed as a single (or in-depth) case study. As we know, there are potential benefits, as well as pitfalls, with single (or small-*n*) case studies (for some examples of such discussions, see Yin 2009; Bennett and Elman 2007; Gerring 2004; and Chapter 8 by Andreas Kruck in this volume). The single case study clearly has limitations, for instance in terms of making generalizations to a larger population of cases. However, viewing a case study as 'an intensive study of a single unit for the purpose of understanding a larger class of (similar) units', it also becomes clear that small-*n* or single case studies may contribute to the way in which we understand and study other, similar cases (Gerring 2004: 342). The study of the Swedish MFA contract can thus be seen as an opportunity to make an empirical contribution and to generate new knowledge about a particular case which, at the time, was poorly understood and clearly under-researched. This in itself is an important undertaking – not least in research on private security where so much of the literature (and thus our knowledge and understanding) has been based on studies of two cases: the US and the

UK. The case also offered a chance to apply theoretical ideas and concepts that have been used in other cases, and thus contribute to the development of a useful framework for analysing some of the key aspects of private security expertise.

In essence, the MFA study sought to make sense of the ways in which the key actors involved in the contracting process *communicate, articulate and make sense of private security expertise*. As noted above, the idea was to see how PSCs draw upon and emphasize different identity traits, but also to study how these images were received and interpreted by state representatives. Empirically, this meant that the study needed to focus on different ‘settings’ in which specialized security skills and knowledge were discussed and articulated, and thus be able to capture nuances and differences in how PSCs construct (and are constructed as) security experts and in the process to identify how and when aspects of military, business and humanitarian cultures or identities come into play.

Of course, there are many settings or contexts in which PSCs articulate and explain their services, skills and specialized knowledge, and these may be fairly distinct or overlapping. For analytical reasons, however, thinking in terms of different settings makes sense because, as we know, a private security company may well draw on very different traits when presenting itself on its website compared with, say, to other commercial actors, to academics, or to international humanitarian organizations. The point in this study was not to cover all of these different sites, but rather to select a small number of manageable and relevant settings that could be used to flesh out differences and similarities. Eventually, two broad settings were outlined: the first was the ‘public’ image, i.e. the official representation of a security company, aimed at prospective clients, the press and the general public. This image is typically found in information produced by the company and published in written form on websites, in booklets or other marketing material. The second setting was the ‘professional’ image, i.e. the image that companies seek to convey when bidding for a contract or communicating with other security professionals. Empirically, the focus in the first setting was on the company (Vesper Group) that was eventually awarded the contract (and had thus been able to convince a previously sceptical MFA of its ability to deliver professional security services). In the second setting, the focus was on the contracting process itself, where two companies (Vesper Group and Scandinavian Risk Solutions, SRS) submitted bids and where MFA security officials reviewed and evaluated the ability of each company to meet the requirements of the contract.

Having come thus far, one of the main challenges for this study was the question of access to data. As students and researchers in this field know, access to reliable data on the activities of private security contractors has often been in short supply – not least when it comes to the specific contents of contracts (see for example Kinsey 2006: 6–7). The security industry is secretive by nature, and the lack of information is particularly pronounced in relation to private security activities in and around violent conflicts or in ‘high-risk’ areas. This situation is very different from studying domestic security privatization in countries such as Sweden, the US or the UK, where it is often much easier to get access to official documents and people to interview, as well as to employ other data collection

methods such as participant observation or surveys. In the context of overseas security contracting, much of the information on private security contracts is classified and thus not easily available for researchers. Finally, in many quarters, the issue of public use of private security in conflict settings remains highly controversial and politically sensitive, making the task of collecting reliable data even more daunting. Trying to overcome these problems, the researcher needs to be aware of (and reflect on) these challenges and the limitations they impose on research, and to be careful not to read selectively, as well as to design studies that avoid (unnecessary and unwanted) bias as far as possible.

At this point, it is useful to consider the nature of documents in general, and the usefulness of combining or mixing document analysis with other sources of data. As Bryman (2008: 526–527) reminds us, there are obvious risks with treating documents (e.g. from the state) as comprehensive or transparent representations of the ‘reality’ of a particular organization. Documents such as reports, minutes of meetings, or indeed invitations to tender for contracts, are created in a specific context, for a specific purpose and with a particular audience in mind and do not necessarily lend themselves to accurate and useful analyses of what goes on in an organization. Thus, ‘if the researcher wishes to employ documents as a means of understanding aspects of an organization and its operations, it is likely that he or she will need to buttress analysis of documents with other sources of data’ (ibid.: 527). Certainly, documents or websites may very well be used as the sole source of data if, for example, the aim of the research is to understand how a larger sample of security companies discursively construct professional and gendered identities and how they frame these images to communicate with prospective clients (for a good example of how PSC websites can be used in this way, see Joachim and Schneiker 2012). The point here is that for an exploratory study aimed at an in-depth understanding of a single case or contract, the research design has much to gain in terms of quality, empirical depth and credibility by mixing different methods for data collection.

In the case of the MFA contract study, the idea was to collect data from the two main settings outlined above. This meant on the one hand to collect data on the ‘public’ image of Vesper Group. As the focus was on the *articulation* of security expertise and knowledge, texts produced by the company for publication on its website served as a good point of departure. However, there is only so much you can glean from these texts, as they are often phrased in fairly general terms and aimed at a very broad audience. For instance, PSC websites rarely comment on specific contracts or their relationship with other public or private security professionals. To complement these data, the company was also approached via email and asked for an interview with a management representative. Using data collection through qualitative interviewing allowed me to move beyond the general and sometimes superficial statements on the website, and to probe a range of issues and questions that were not covered in the online material (such as their view of the general debate on the private security industry, or their take on the Swedish market for PSCs). This provided a fuller picture of the self-image of the company and improved the prospect of addressing the research question.

In the second setting, a similar line of thinking led to a mixed data collection approach. The idea was that official documents from the Swedish state relating to the contracting process, i.e. the invitation to tender, the bids or proposals and the subsequent MFA evaluations and decisions, would make it possible to see what services the MFA specified in the call, what criteria they set up to evaluate the security expertise of private security providers, as well as the ways in which the companies strived to meet these criteria and in the process describe and explain their competence in different areas. In addition, the MFA evaluations and decisions would most likely say something about how companies' skills and expertise were deemed relevant, useful, legitimate and persuasive by state security officials. However, and as discussed below, the Swedish official state documents turned out to be of questionable quality, and complementing these sources of data became all the more important. In this case, making a credible and meaningful analysis of private security expertise and the 'professional image' of private security actors in this contracting process would not have been possible without the data from an interview with MFA security officials. Before elaborating on the application of different methods for data collection and the problems faced during this phase, a few words on data analysis are in order.

Departing from the question about the construction of private security expertise and the theoretical framework derived from previous studies on PSC identity and self-images, the analysis drew primarily on a set of methods for analysis that are sometimes referred to simply as "thematic analysis." There are many ways in which such an analysis can be made, and there is no "standard" formula or procedure. Broadly speaking, it is a flexible qualitative approach focusing on eliciting key themes and meanings from mainly (but not necessarily only) textual sources, either in order to inductively explore a phenomenon or to test hypotheses (Guest *et al.* 2012). Or, as Braun and Clarke write, thematic analysis 'examines the ways in which events, realities, meanings, experiences and so on are the effects of a range of discourses operating within society' (Braun and Clarke 2006: 81). Thus, the approach is less quantitatively oriented than classic content analysis; it goes 'beyond counting explicit words or phrases and focus on identifying and describing both implicit and explicit ideas within the data' (Guest *et al.* 2012: 10). Thematic analysis essentially focuses on identifying patterns or meanings in and across qualitative data sets. The analysis of the MFA contract was partly deductive in that it was guided by pre-defined 'main themes', i.e. theoretical categories and concepts collated from previous research on private security expertise that help give direction to the analysis. These were mainly drawn from Leander's work on security knowledge/expertise and Carmola's three organizational types or cultures. This meant that the 'military', 'humanitarian' and 'business' dimensions of PSC identity were treated as main themes that we can expect to find in different settings where their expertise is described or explained. In the analysis, the main task was to find out in what way and to what extent the main themes appeared in the case of the Swedish MFA contract, what weight they were given, as well as to identify new supporting or sub-themes (such as the theme of 'Swedishness' discussed below).

To sum up, the study was designed as an in-depth single case study of the 2008 MFA contract, relying on a design using mixed methods for data collection to provide a basis for a 'thematic' analysis of private security expertise in two different settings. As with all research designs, this also comes with its particular benefits and challenges. One of the strengths of the design is arguably in the combination of different sources of qualitative data. Combining data from official state and private sources with semi-structured interviews clearly yielded a much richer image of private security expertise than could have been gained from either of the two types of sources. The interviews also allowed me to probe some of the themes and topics found in the documents further, and to cover issues not found in the documents. The combination increased both the quality and credibility of the analysis, and made it possible to overcome some of the problems associated with relying on a narrow selection of potentially biased sources. However, the design of the study also has clear limitations, especially in terms of empirical scope and representativeness of results but also, as we shall see, in terms of quality of data.

Data collection: documents and interviews

As noted above, one key challenge for this study was the lack of previous research on the Swedish use of private security in high risk or conflict areas. Thus, one of the first tasks was to try to find out more about the contract and its contents. Apart from a few questions in the Swedish Parliament and some short articles in newspapers, the discussion of the Swedish use of PSCs in high-risk settings was virtually non-existent. The idea was the official documents from the MFA could provide some basic information, as well as insight into the 'professional image' of the companies bidding for the contract. Thus, on 25 May 2010, a Freedom of Information request was sent via email to the Ministry for Foreign Affairs, asking them to release, for research purposes, all documents pertaining to the 2008 contract. When the official reply arrived on 18 June, it became evident that this source had some serious limitations. The documents (around 100 pages in all) did provide a fairly clear image of the contracting process, excerpts from the bids from the two PSCs along with the MFA evaluations, as well as some details about the nature, extent and cost of the contract. However, many sections or parts of sections were partly or wholly blanked out, formally with reference to national security interests and the Public Access to Information and Secrecy Act. The blanked out parts were typically in sections on key assignments and requirements, arrangements and organization of service provision at the embassies and use of equipment (including weapons). While the need for confidentiality in some areas can be readily understood, the perceived need for secrecy on so many issues nonetheless came as a surprise. At the same time, the blanked out sections (amounting to around 30–40 per cent of the total number of pages released) represented a result in themselves; they suggested that even in a fairly open democracy such as Sweden, many issues surrounding the contracting for security services abroad is seen as sensitive.

Methodologically, it would have been very difficult to rely on the documents alone, as they were clearly far from complete and conveyed at best a partial and distorted image of the contracting process. The formal decision to suppress certain details also included instructions on how to appeal. If an appeal was sent in, the Freedom of Information request would be reviewed by a government minister. Submitting an appeal might have resulted in more documents being released, but it could have diminished the chances of securing an interview with MFA security officials. In this case, no appeal was made; instead, the MFA Secretariat for Security were contacted and asked to participate in an interview. With hindsight, an appeal would probably not have affected the chances of securing the interview, but at the time it appeared less risky to ask for an interview without an appeal and thus try to rectify some of the shortcomings of the documents.² As will be discussed below, despite the limitations of this source of data, the documents did provide basic information about the contracting process that proved fruitful both for the construction of interview questions, and for an initial analysis of key themes related to private security expertise.

Regarding the other source of documents, this time collected from the PSC website, the experience was similar. The texts on these pages are obviously written for some particular purpose(s) and intended for a broad audience. Thus, while the texts describing the company's services, expertise, ethical code of conduct or the professional background and experience of staff do reveal a lot about how it wishes to construct its public image, it does not provide us with the full picture. Although the omissions and gaps on PSC websites are not as visible as in the documents released by the MFA, they are clearly there, and we must again ask ourselves if these sources are enough for our scientific purposes or if they need to be complemented with other sources. In the case of the MFA contract study, the very obvious limitations of the official state documents along with the feeling of not being able to get at the 'public image' of private security expertise by analysing only online material were clear indications that something more was needed. In this particular study, there were few options besides conducting research interviews with both state and PSC officials.

As many students of private security have experienced, interviewing PSC managers or state security officials requires a fair amount of planning, practice and, last but not least, luck. These are busy people, not always easy to get hold of, and occasionally suspicious of academics asking critical questions about (what are often perceived as) sensitive issues.³ The first challenge is to find the right persons to interview and to get them to participate. This normally requires a formal letter with a description of the research project and the purpose of the interview. In my experience, PSC managers and state security officials are generally willing to take part in research projects as long as they can be convinced of the usefulness of participating. For PSC managers, this usefulness might be found in the chance to explain and motivate their activities, or to get their message across. Of course, this presents researchers with something of a dilemma: we want to gain access to people on the 'inside' to learn about their organizations, activities, motivations or self-images but at the same time we do

not want to simply (and uncritically) reproduce their ideas and world-views. On the other hand, we do not want to be seen as overly critical or biased, in which case company managers would most likely refuse to take part.

The key here is, as Kinsey has also pointed out, to retain 'a certain degree of open mindedness about the [private security] industry' (Kinsey 2006: 7). Emphatically, this does *not* mean accepting the world-views of the security industry or refraining from asking critical questions. It does mean, however, that when approaching PSCs (or, for that matter, state security officials), we need to build rapport, to come across as genuinely interested in learning from their experiences and as capable of (re-)presenting their stories in a professional manner. In the MFA contract study, as well as in previous and later work, these ideas have been translated into a formula for securing and conducting interviews that has generally worked well. In the first contact with prospective interviewees, I have presented the research project in general terms and the purpose and format of the interview in more detail. At this stage, and following the ethical guidelines of the Swedish Research Council, prospective interviewees are also provided with information about consent, the right to opt-out, anonymity and how interview data will be treated once the interview has been conducted. This means that interviewees retain a fair amount of control over their stories. A verbatim transcript of the (usually recorded) interview is sent to them afterwards, and interviewees are given a chance to comment on quotes or references that will be used in the published study. As several of my interviewees have told me, the feeling of being in control has been one important reason for them to accept participation.⁴ While one could argue that too much control is handed over to interviewees, I think that this is ethically sound practice that serves to protect the people we involve in our projects. In my experience, it has not affected my ability to ask critical questions, or to make independent interpretations of interview data.

In the MFA contract study, two semi-structured interviews were conducted: one in August 2010 with the head and deputy head of what was then the Secretariat for Security at the Swedish MFA; the other in October 2010 with a representative of Vesper Group management. The rationale for approaching these (and not other) individuals was that they were the ones closest to the contracting and decision-making process, and could thus be expected to have unique insight and knowledge that could not be found elsewhere. Additional interviews with PSC employees and/or MFA staff could have generated more data. In this case, however, interviewing three of the most centrally placed persons in the case yielded enough insight into the contracting process and served as a sufficient complement to the data collated from the documents and the website. Both interviews were organized around a list of main questions and topics, but also with enough flexibility to follow up on interesting leads. The research question as well as my initial readings and analyses of the document and website data served as a basis for designing the interview guides and main topics. The opening questions revolved around the debate on security privatization, and the purpose was to get interviewees to talk about this topic in general or neutral terms, but also to get them to position themselves vis-à-vis common views and statements

(e.g. ‘What is your take on the common understanding of PSCs as mercenaries?’ or ‘What were your reactions to the Nisour Square shootings in 2007?’). The following themes covered the background to and details of the MFA contract (e.g. ‘What services do you think private security contractors can and should provide in this case?’ or ‘What is it that your company brings to the contract with the Swedish MFA?’). Apart from providing insight into the contracting procedure, these themes also opened up opportunities for questions about private security expertise and professionalism. The idea was to ask similar questions to both MFA officials and PSC representatives, and to get them to talk about what it is that makes private security providers useful, acceptable, and legitimate actors. This involved asking questions about how they (Swedish PSCs in general or Vesper Group in particular) are similar to, or different from, other actors such as state militaries or police forces, or indeed other private security actors. The third and final main theme covered questions of regulation, control and oversight. Here, the idea was to discuss issues of private security regulation and state control in general as well as in the specific case. Again, the aim was to learn more about the specifics of the particular contract, but also to get interviewees to reflect on issues such as the use of force by private security actors, the need for state control and oversight of these actors, the differences between private and state security actors in terms of regulation and control, and on the private security industry more generally.

Each interview lasted about one hour, and the full transcripts yielded around 40 pages of text for further analysis. The interviews provided both basic information and, most importantly, ample reflection on issues related to the contract and the issue of private security expertise. Of course, some details were impossible to discuss during the interviews because they were classified or too sensitive. On these occasions, common interview techniques such as lifting the discussion to a more general and non-threatening level or playing devil’s advocate proved very useful, avoiding the ‘out of bounds’ or evidently uncomfortable areas of the MFA contract while at the same time getting interviewees to reflect more on relevant topics. All in all, the interview data complemented the other sources well and provided a useful empirical basis for the subsequent analysis.

Data analysis: identifying themes and meanings related to private security expertise

As indicated above, the analysis of the empirical data was deductive in the sense that it was guided by a theoretical framework where main themes and concepts such as ‘private security expertise/knowledge’ and the ‘military’, ‘humanitarian’ and ‘business’ dimensions of PSC identity were drawn from previous research. The purpose of the analysis was to see if and how these themes appeared in the context of the MFA contract, how they were bestowed with meaning by key actors and in the two different settings, as well as to identify new and relevant sub-themes and concepts. At this point, the online material, along with the official state documents received through the Freedom of Information request and

the data obtained through the two interviews made up the bulk of the empirical data for the study (all in all about 200 pages of text). With hindsight, these sources provided a sufficient empirical basis for the exploratory purpose of the study, but ideally a more comprehensive and sophisticated analysis could have been done with more data, for instance in terms of documents on internal MFA discussions leading up to the decision to contract for security, or interviews with a larger sample of PSC and MFA or government representatives and employees.

Returning to the analysis of the data, it did not strictly follow any of the many existing formulas for ‘thematic analysis’, such as the six steps suggested by Braun and Clarke (2006). It did, however, proceed in three main steps or phases that served the purpose of systematically organizing the data and identifying recurring patterns and meanings. The first step was a close reading of the online material and the official state documents. This part of the analysis was both a kind of data familiarization, but also functioned as a way to find gaps and questions to be developed in the interviews. The second step was the actual coding of all the data once the interviews had been conducted. In practice, this meant manually identifying and marking interesting pieces of the data (key words, phrases, expressions) with a particular focus on questions and patterns relating to the articulation and reception of PSC self-images. Since the datasets were relatively small, the coding was done manually, using different colours to mark different themes and sub-themes. The coding generated a large number of text extracts that could then be grouped together and analysed in relation to the main themes of security expertise and the military, humanitarian and business aspects of private security identity. The final step involved making visual (thematic) maps where the main themes were connected to relevant and supporting sub-themes and comparing the different datasets (i.e. the online material, the official documents and the interview transcripts) to see how pre-defined and new themes lined up in relation to the ‘public’ and ‘professional’ settings.

In the published article, the main body of the text was structured around the two settings and the ways in which pre-defined and new themes were employed and given different meanings. In order to lend transparency and credibility to the analysis and to show how private security expertise was constructed in this case and how it was related to business, military and humanitarian qualities, a number of quotes from the different sources were used. In terms of results, it was clear that all of the main themes were present in the data, but to different degrees and with different meanings. For instance, references to a ‘military culture or identity’ were kept to a minimum in one setting while being a key theme in another. This clearly lends support to the idea of PSCs as ‘hybrid’ or ‘protean’ actors and as clearly being ‘more than one thing at once’ (Carmola 2010). It also revealed some very interesting sub-themes and concepts. For instance, it showed how the idea or theme of ‘Swedishness’ appeared in both the PSC and MFA interviews, and was used to motivate and explain security expertise, but also to construct a ‘humanitarian’ self-image. It also showed how distinctions between ‘public and private security actors and expertise’ tended to become distorted and mixed up, especially when state security officials sought to rationalize the decision to use

private contractors. Taken together, the study showed how images and understandings of private security companies as ‘security experts’ are complex, and that questions of context and audience are key if we want to understand how these images are constructed, received and legitimised.

Discussion

The study discussed in this chapter employed an exploratory approach to address a research question about the formation of private security expertise in different settings. As discussed above, the combination of different sources of data and methods for data collection proved fruitful as it helped overcome some of the difficulties of gaining access to information on private security contracting. Complementing the partial and in this case highly distorted image provided by official documents was necessary in order to improve the quality of data and thus the credibility of the study, but also to facilitate comparisons. However, the data collection process was certainly not without challenges.

In research on non-state actors in international security, conducting semi-structured interviews can be very useful, as they certainly were in this case, although finding interviewees who are willing to speak ‘on the record’ can be both challenging and time consuming. Interviewing also raises some fairly thorny questions about how to approach prospective interviewees, how to get them to move away from the ‘official’ or public relations rhetoric, as well as how to deal with ethical issues related to participation and use of interview material. In this context, we should be careful not to overestimate the possibility of gaining an ‘inside’ perspective through semi-structured interviews. To really move beyond official narratives, we need to consider alternative methods, such as participant observation and field work over an extended period of time (see Chapter 12 by Tessa Diphoorn in this volume). In addition, the themes and issues that you can reasonably expect to cover in interviews are limited – especially if the aim is to find out about details of contracts and PSC activities. Private security representatives and state security officials are often very cautious and sometimes wary of researchers and will not necessarily be open to discussions of what they feel are sensitive issues.

In this case, however, sufficient data were collected to allow for an analysis of the construction of PSC security expertise in the context of the Swedish MFA contract. The analysis provided interesting and novel insight into the ways in which PSCs are constructed and legitimized as security experts, suggesting some ways forward for future studies, preferably with more cases and access to more data. Methodologically, this study could not have been conducted by focusing on official documents alone. Even in a relatively open democratic society, the issue of private security contracting is highly sensitive, and access to reliable data can be very restricted. With hindsight, conducting more interviews would probably have been useful to overcome the deficiencies of the official documents. Yet it is also important to underscore that it is not necessarily the number of interviews that determines the quality of data; in this case, the central placement of the three

persons interviewed sufficed to lend considerable added value to the analysis. Ideally, however, the exploratory study of the MFA contract should be followed up, perhaps with a comparative case study design. This would allow for a more 'saturated' or comprehensive analysis, making it possible to draw more far-reaching and general conclusions. Finally, the fairly rudimentary thematic analysis conducted in this study could be developed or – depending on data availability, the theoretical approach and formulation of research questions – exchanged for more technically advanced methods.

On a more general note, the experiences from this study show that the combination of different sources of qualitative data collected from different 'settings' allowed the analysis to move at least some distance beyond official accounts, to probe central themes more deeply, and thus to better capture the complexities and ambiguities of private security expertise. In a wider perspective, conducting research on non-state actors in international security where access to data is limited or where data can be suspected of being heavily biased, this kind of qualitative 'data triangulation' might serve as a useful methodological strategy.

Notes

- 1 One study published around this time was a report by Andreas Bergman for the so-called PRIV-WAR project, focusing on the legal status of PSCs in different countries, including Sweden (Bergman 2010). In the report, Bergman found that private security companies operating abroad are not specifically covered by Swedish law, and also that statements from Swedish state officials (including the Ministers of Justice and Foreign Affairs) from 2007 onwards appeared to suggest that no such regulation was needed.
- 2 In 2013, the MFA contracts were up for renegotiation and a new Freedom of Information request filed for another study. Again, many sections were blanked out. This time, an appeal was made and the original request was put before a government minister. In the decision (signed by the then Minister for International Development Cooperation, Gunilla Carlsson), the government decided to provide some additional information (a few paragraphs in over 200 pages) but also decided to turn down the request once more. Motivating its decision, the government notes that even though the information is to be used for research (a condition under which restrictions can be lifted), and even though the government can make information for such purposes available under conditions of confidentiality, the 'risk of damage' to Swedish national security interests and business relations could not be satisfactorily avoided. This experience underscored the sensitivity of the issue in the eyes of Swedish state officials, the silence that surrounds security contracting, as well as the difficulties of getting access to reliable (and comprehensive) data on state-PSC contracts – at least through official documents alone (for further discussion of these later experiences, see for example Berndtsson and Stern 2013: 70–71).
- 3 PSC managers and state security officials can be seen as representatives of different 'elites' and studying them comes with specific methodological challenges (for some discussions of these issues, see for example Moyser and Wagstaffe 1987; Hertz and Imber 1995).
- 4 There are many ways in which to plan and conduct semi-structured interviews (for some general discussions, see Bernard 2000: 190–207, Bryman 2008: Chapter 18). For a more in-depth discussion of conducting semi-structured interviews in private security research, see Chapter 5 by Anja Mihr in this volume. In addition, there are many useful texts about specific techniques and discussions about the ethical issues of qualitative interviewing (Rubin and Rubin 2005; Bryman 2008: Chapter 5; Seidman 2006).

References

- Abrahamsen, Rita, and Michael C. Williams. 2011. *Security beyond the State: Private Security in International Politics*. Cambridge: Cambridge University Press.
- Bennett, Andrew, and Colin Elman. 2007. 'Case Study Methods in the International Relations Subfield'. *Comparative Political Studies* 40(2): 170–195.
- Bergman, Andreas. 2010. *The Regulation of Private Military and Security Services in Sweden*. Florence: European University Institute.
- Bernard, H. Russell. 2000. *Social Research Methods: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage Publications.
- Berndtsson, Joakim. 2012. 'Security Professionals for Hire: Exploring the Many Faces of Private Security Expertise'. *Millennium: Journal of International Studies* 40(2): 303–320.
- Berndtsson, Joakim. 2013. 'Exploring PSC-military relations: Swedish officers and the private security sector in peace operations'. *Cooperation and Conflict* 48(4): 484–502.
- Berndtsson, Joakim, and Maria Stern. 2011. 'Private Security and the Public-Private Divide: Contested Lines of Distinction and Modes of Governance in the Stockholm-Arlanda Security Assemblage'. *International Political Sociology* 5(4): 408–425.
- Berndtsson, Joakim, and Maria Stern. 2013. 'Sweden: Public Servants from the Private Sector', pp. 58–78 in *Commercialising Security in Europe: Political Consequences for Peace Operations*, A. Leander, ed. London and New York: Routledge.
- Braun, Virginia, and Victoria Clarke. 2006. 'Using Thematic Analysis in Psychology'. *Qualitative Research in Psychology* 3(2):77–101.
- Bryman, Alan. 2008. *Social Research Methods*. Oxford: Oxford University Press.
- Carmola, Kateri. 2010. *Private Security Contractors and New Wars: Risk, Law, and Ethics*. London: Routledge.
- Gerring, John. 2004. "What is a Case Study and What is it Good for?" *American Political Science Review* 98(2): 341–354.
- Guest, Greg, Kathleen M. MacQueen, and Emily E. Namey. 2012. *Applied Thematic Analysis*. Los Angeles, CA: Sage Publications.
- Hertz, Rosanna, and Jonathan B. Imber. 1995. *Studying Elites Using Qualitative Methods*. Thousand Oaks, CA: Sage Publications.
- Joachim, Jutta, and Andrea Schneiker. 2012. 'Of "True Professionals" and "Ethical Hero Warriors": A Gender-Discourse Analysis of Private Military and Security Companies'. *Security Dialogue* 43(6): 495–512.
- Kinsey, Christopher. 2006. *Corporate Soldiers and International Security: The Rise of Private Military Companies*. New York: Routledge.
- Krahmann, Elke. 2011. 'Beck and Beyond: Selling Security in the World Risk Society'. *Review of International Studies* 37(1): 349–372.
- Leander, Anna. 2006. 'Privatizing the Politics of Protection', pp. 19–33 in *The Politics of Protection: Sites of Insecurity and Political Agency*, J. Huysmans, A. Dobson and R. Prokhovnik, eds. London: Routledge.
- Leander, Anna. 2007. 'Regulating the Role of Private Military Companies in Shaping Security and Politics', pp. 49–64 in *From Mercenaries to Market: the Rise and Regulation of Private Military Companies*, S. Chesterman and C. Lehnardt, eds. Oxford: Oxford University Press.
- Moyser, George, and Margaret Wagstaffe. 1987. *Research Methods for Elite Studies*. London: Boston: Allen & Unwin.

- Rubin, Herbert J., and Irene Rubin. 2005. *Qualitative Interviewing: The Art of Hearing Data*. 2nd edn. Thousand Oaks, CA: Sage Publications.
- Seidman, Irving. 2006. *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. New York: Teachers College Press.
- Yin, Robert K. 2009. *Case Study Research: Design and Methods*. Los Angeles, CA: Sage Publications.

7 Discussion chapter

Comments on ‘interpreting texts’

Jutta Joachim

Which method?

Which method to choose is not an easy question to answer. Unless there is an *a priori* preference on the part of the researcher for either qualitative or quantitative methods, what one seeks to accomplish is decisive in terms of a response. In the case of the chapters in this section, the goal is the interpretation of texts broadly defined. But even if the ‘what for question’ has been settled, the authors demonstrate that there is a plethora of methods to choose from. While the methods presented here can loosely be subsumed under the umbrella of content analysis, intended to capture how events or actors are perceived and are made sense of, they highlight very vividly that this type of method comprises various strands, each of which requires different tools or research steps and serves different purposes.

Alexander Spencer (see Chapter 3) introduces *narrative analysis*, which in his eyes ‘helps illustrate the dominance of certain stories in public discourse while at the same time indicating the void left by stories that remain untold or marginal’ (Spencer, this book). In his case, narrative analysis makes apparent how rebels during the 2011 revolution in Libya were romanticized by both British media as well as political elites, while human rights violations perpetrated by them received close to no attention (Spencer, this book). Joakim Berndtsson (see Chapter 6, this book) by comparison, employs *thematic analysis* to identify ‘key themes and concepts [...] examine how they were used in the construction of private security [professionals]’ and ‘how they become legitimate non-state security actors’. Contrary to narrative analysis, which is based on certain preconceptions of how narratives are constructed and for what purpose, thematic analysis as presented by Joakim Berndtsson (this book) is rather explorative in nature aimed at ‘generating new knowledge’. It is suited for mapping the range of possible manifestations of what one is looking for or, as Joakim Berndtsson puts it, ‘identifying patterns or meanings in and across qualitative data sets’.

Anja Mihr (see Chapter 5) discusses *semi-structured interviews* as an interpretative method and a means to elicit ‘why’ or ‘how something happened’, and why ‘organizations and institutions do what they do’ as well as their motivation

behind it. Finally, Magnus Dau and Marlen Martin (see Chapter 4) present a mixed-methods-design combining *quantitative computerized sentiment analysis* with *qualitative content analysis*. While the former allows one to classify and group large amounts of text – in their case tweets – according to whether the message can be characterized as positive, negative or neutral, content analysis helps to dissect what the dominant positive and negative frames are.

The methods introduced here have strengths and limitations that come to the fore in a comparative perspective which I will adopt, organizing my comments around the different research stages beginning with data collection and continuing with the interpretative analysis.

What constitutes text and how to gain access to it

The chapters in this volume lend force to the assertion that '[a]nything that is intended to communicate a message is usable as material for content analysis' (Hermann 2008: 152). What connotes 'text' varies widely ranging from official (state) documents (see Chapter 6 by Joakim Berndtsson and Chapter 3 by Alexander Spencer), to newspaper articles (see Chapter 3 by Alexander Spencer), words uttered during interviews (see Chapter 5 by Anja Mihr and Chapter 6 by Joakim Berndtsson), but also homepages (see Chapter 6 by Joakim Berndtsson) or twitter messages (see Chapter 4 by Magnus Dau and Marlen Martin). While these various texts exhibit unique characteristics, their treatment in the different chapters also makes clear that none of them is 'simply' text, but already arrives quite often at the researcher's desk with a layer of interpretation.

All of the texts collected and used exhibit unique characteristics. They lend force to Joakim Berndtsson who, in reference to Alan Brymann (2008: 526–527), and in connection with government documents including 'reports, minutes of meetings, or indeed invitations to tender for contracts', asserts that all 'are created in a specific context, for a specific purpose and with a particular audience in mind and do not necessarily lend themselves to accurate and useful analyses of what goes on in an organization' (Berndtsson, this book). The websites of private security companies, for example, are often used for promotional, sale – as well as image enhancing purposes while tweets may have been written with different aims in mind, be it networking, building one's reputation, or distributing information to name only a few. Given their particular characteristics, rarely do the authors analyse the different types of texts in isolation, but instead in a triangular fashion.

This serves the purpose to, on the one hand, cross-check and solidify one's findings. By pairing newspaper coverage with official statements of British elites during the course of parliamentary debates or Security Council meetings, Alexander Spencer is able to show that the romanticized narrative of the Libyan rebels is not just an invention of the media, but also shared by political elites. Nevertheless, triangulation has a further meaning. According to Norman K. Denzin and Yvonna S. Lincoln, it is not just a tool of validation. Instead and, on the other hand, it is a strategy that adds 'rigor, breadth, complexity, richness and

depth to an inquiry' (Denzin and Lincoln 2011: 5). Seen this way, triangulation allows us to construct both a coherent as well as comprehensive story. It accounts for the imitability of different texts, the intention with which they have been produced and, as Iver Neumann observes, '...the medium through which they are published' (Neumann 2008: 68).

The uniqueness of texts becomes unmistakably clear when reading the chapter by Magnus Dau and Marlen Martin related to twitter messages. Their sheer volume contributes to the exceptionalism of the messages and makes it nearly inevitable to rely on quantitative methods. Furthermore, tweets are distinct from more conventional types of documents because they are, as Magnus Dau and Marlen Martin (Chapter 4, this book) point out, 'arenas for public deliberation and public opinion formation'. They make it possible for researchers not only to capture data in 'real time' and on 'current themes', but also ongoing 'communications' comprised of messages sent and the reactions to them. Hence, reactions to political events, developments, or statements do not have to be actively solicited by researchers by conducting, for example, interviews, surveys or by relying on opinion polls. Instead, they are delivered 'free house' without delay and on the topic. Finally, tweets differ from other texts in yet another respect. Their authors often remain anonymous and provide little information about their true identity, an aspect to which I will return when turning to my reflections on interpretative analysis.

Interviews too reveal interesting particularities. Contrary to the other sources mentioned in the different chapters, which are already given and demand of the researcher solely to choose between them, interviews only come into being through the encounter and communication between interviewee and researcher. In that sense, they have a certain 'unpredictability' as far as the result is concerned. The emerging text is not merely dependent on the researchers' 'social competences' and the quality of their more or less standardized questions, but also hinges on how much or how little the respondent is willing to reveal (Mihr, Chapter 5, this book). Interviews, according to Anja Mihr 'leave room for personal reflections, story-telling and deviations'. While they allow, as Joakim Berndtsson, who also relies on them, points out (Chapter 6, this book), to 'move beyond the general statement' of publicly available material and 'probe a range of issues', conducting them involves a number of challenges that influence what the text will look like in the end. These challenges include, among others, the choice of the interviewees or experts, the relationship the researcher maintains with them, the language skills as well as identity-related factors of both researcher and respondent or the context in which or the means with which the interview is conducted. All of these factors contribute in their own way to what might be characterized as a 'living document'.

The challenges encountered when conducting interviews may suggest that the other types of texts are less cumbersome and rather straightforward when it comes to their retrieval and acquisition. Yet, as the chapters demonstrate, this is not necessarily the case. The official documents of the Swedish Ministry of Foreign Affairs that Joakim Berndtsson employs had neither been easy to get access to – the author and his research team had to apply for them to be released

– nor were they available in their entirety. Prior to their release, the ministry had run a check as to what is suitable for public consumption and what needs to remain secret for national security purposes (Berndtsson, Chapter 6, this book). Although the documents provided ‘a fairly clear image’ of different aspects of the contracting process with private security companies, they contained ‘blanked out parts [which] were typically ... sections on key assignments and requirements, arrangements and organization of service provision at the embassies and use of equipment ...’ (Berndtsson, Chapter 6, this book). Joakim Berndtsson’s account of the data collection phase illustrates once more how texts are unique artefacts, although it points to an additional source that contributes, along with their design, purpose or carrier, to their specific ‘contours’: the filtering processes to which they have been subject.

That the authors in this volume openly share how they went about collecting their texts and the difficulties they were faced with is laudable since these aspects of doing research are rarely covered to such an extent in method books. However, their observations also raise questions that deserve further reflection, especially in light of the fact that rarely is any text just text; it brings with it its own history and carries with it multiple interpretative layers. Can we, for example, combine unproblematically different data sources, such as documents and interviews? If so, what precisely do we learn from some, but not others and what does this variation tell us? Even if different texts tell us a similar story, such as for example in the case of Alexander Spencer or Joakim Berndtsson, what do the data sources as such in addition to the content of the texts reveal about how narratives or stories and dominant themes are (re-)produced, by whom and in what way? Does it make a difference whether it is a newspaper that (re-)produces the narrative or whether statements by governmental officials or documents of political parties or civil society organizations do, whether a tabloid carries the story or whether it is embedded in a report of a human rights organization?

Answering questions of this kind can contribute to a richer interpretation, yet also involves a shift in perspective. Rather than treating texts as a means to an end, i.e. to get at the story, and merely a data source, it requires for us to conceive of the text as such as endogenous to our interpretative analysis and as an important piece to our puzzle. Depending on the source, we then may assign different weight to the narrative and perhaps conceive of it as more or less critical, as more or less status quo enhancing, or we may attribute more or less transformative potential to it.

From text to text: interpretative analysis

Interpreting text is sometimes dismissed as a subjective and arbitrary enterprise that hinges on the researcher and what he or she deems important. Yet, as the chapters in this section demonstrate and as Donald Moon (1975: 173) observed:

[t]he most obvious standards are coherence and scope: an interpretative account should provide maximal coherence or intelligibility to a set of social

practices, and an interpretative account of a particular set of practices should be consistent with other practices or traditions of the society.

Each of the methods used here entails a particular strategy to accomplish this.

Of the methods employed, narrative analysis seems to be the most straightforward in this respect. This might appear as somewhat counter-intuitive at first glance. Narratives are particularly ‘contextually thick’ (Patterson and Monroe 1998: 316) and provide ‘a rich source of information about how people make sense of their lives, about how they construct disparate facts and weave them together cognitively to make sense of reality’ (Patterson and Monroe 1998: 315). Yet, it is precisely this ‘thickness’ and richness, that affords a more rigorous ‘road-map’ as to what to look for when analysing narratives, specified in Alexander Spencer’s case by (1) the setting, (2) the characterization of the actor(s), and (3) the causal emplotment.

Similar to narrative analysis, thematic analysis as employed by Joakim Berndtsson in this book is far from random. Aimed at uncovering key themes in relation to private security companies, the author constructs a coding system deductively relying on ‘pre-defined “main themes”, i.e. theoretical categories and concepts collated from previous research on private security expertise’. While the themes and patterns prevalent in the literature provided orientation as far as content is concerned, drawing on the work of Virginia Braun and Victoria Clarke (2006) Joakim Berndtsson also had guide posts for how to proceed and for ‘...systematically organizing the data and identifying patterns and meanings’. Subsequent to a close read of online and official state documents, he first manually coded the texts before grouping the marked fragments in relation to the dominant themes. He finished by constructing visual maps to determine dominant as well as sub-themes, the way in which the themes that evolved from the materials related and corresponded to each other as well as to different settings. Finally, sentiment analysis – probably the most demanding in terms of preparation – also follows a protocol. It involves an iterative process between software and texts to teach the computer what constitute (very) positive, (very) negative, or neutral words so that it can assign the appropriate sentiment autonomously to thousands of twitter messages.

While having a script of how to proceed might ensure a coherent interpretation, it also involves trade-offs. In particular, it may suggest a stringency or linearity that narratives or other textual materials rarely exhibit and it may draw attention away from the more intangible fabric, that is, the myths, the symbols or historical experiences on which stories rest and that lend power to them. Why does, for example, the exotic rebel play such an important role and how is that narrative informed or moulded by a post-colonial or geo-political script? Capturing this fabric is perhaps what distinguishes narrative and thematic analysis from discourse analysis, which also is interested in ‘dominant representations of reality and one or more alternative representations’ but also their deeper meaning (Neumann 2008: 70). Compared with content analysis, which is interested in the story as such and the kind of story that is told, discourse analysis makes us dig

deeper and move beneath the surface. It makes us look back in time and ask what certain symbols or tropes represent or mean within a particular historical societal context or what contradictions and tensions they reveal.

Contributing to the consistency and coherence of the analysis, the preoccupation with looking for certain things may prevent one from seeing (or hearing in the case of interviews) what is also there, but does not fit into the pre-given categories. Such aspects may include, for example, a gendered script that contributes in the narrative about rebels, and the construction of white Western masculinity as hegemonic and that of Southern coloured men as different, other and subordinate or, in the case of private security companies, to the privileging of professional and military masculinity.

Similar to road-maps, contextual knowledge, or what Iver B. Neumann in relation to discourse analysis refers to as 'a basic level of cultural competence to recognize the shared understandings that create a common frame of reference, which makes it possible for people to act in relation to one another' (Neumann 2008: 64), can be argued as being essential for a coherent account. Contextual knowledge allows us to identify and capture what sticks out, what is dominant, or what is left out, and to uncover underlying power asymmetries, recurring patterns or themes. As Molly Patterson and Kristen Monroe (1998: 320) note, '[t]he stories [and hence the narratives] we tell are profoundly influenced by what is possible and what is valued within our culture'. Ironically perhaps, it is the quantitative sentiment analysis that makes this utterly clear.

What constitutes positive, negative, or neutral is far from obvious and differs across context(s) broadly. As Bo Pang and Lillian Lee with the exclamation 'go read the book' illustrate, 'the exact same expression can indicate different sentiment in different domains' (Pang and Lee 2008: 21). While it would be indicative of a positive sentiment when used in a book review, it most certainly is negative when it is the résumé of a film review (*ibid.*). If we understand context in a yet narrower sense and apply it to the order in which words appear, it becomes even more apparent how difficult it can be to determine the overall sentiment of text segments (Dau and Martin, Chapter 4, this book). Take again an example by Bo Pang and Lillian Lee of a movie review: 'This film should be *brilliant*. It sounds like a *great* plot, the actors are *first grade*, and the supporting cast is *good* as well, and Stallone is attempting to deliver a *good* performance. However, it can't hold up' (Pang and Lee 2008: 22, emphasis added). The number of positive words would lead to an overall positive sentiment, when in fact the movie is ultimately judged negative (see also Chapter 4 by Magnus Dau and Marlen Martin, 4). In addition, the nature of texts themselves may be suggestive of a certain context. As Magnus Dau and Marlen Martin note, tweets, which, even if comprised of relatively few words, are 'idiosyncratic and dominated by colloquial language', are different from the formal language the software is based on and therefore require an active act of translation on the part of the researcher.

While fundamental for interpretation, there are externally imposed limits of what we can know. Several of the chapters allude to such constraints, but do not

elaborate on their implications. As already noted above in connection with the sentiment analysis of tweets, information about the users is often difficult to acquire, such as their race, colour, gender or their geographical location. Many messages are posted anonymously and can even be computer-generated or users may use pseudonyms. Although this ‘veil of ignorance’ might be deemed as an advantage since it allows us to interpret text on its own merits, it is, however, the identity of the interlocutor that lends meaning and provides context to what is uttered. Consider the statement on the homepage of a private military and security company that claims to value diversity. The statement as such suggests a positive sentiment. However, when analysed in combination with an accompanying photo, which shows a white trainer in front of coloured people looking down, the statement takes on a different meaning. Rather than of inclusion, it speaks of existing power asymmetries and would lead us to think of the alleged diversity rather sceptically.

Furthermore, and in contrast to narrative analysis, a computer is only sensitive and trained to evaluate a given text, but unable to capture the invisible (con-)text in the sense of what is not said. It cannot detect whether tweets are posted in an overwhelming fashion by individual users and might therefore be indicative of existing hierarchies, nor is it able, as Bo Pang and Lillian Lee point out, to expose seasoned writers who can cloak negative sentiments through positive words (Pang and Lee 2008: 21). While the limits to what we can know about context are perhaps more prevalent when analysing social media, the chapters in this section make apparent that we may also encounter them when analysing more conventional types of texts.

The chapter by Joakim Berndtsson on private security companies draws attention to the fact that political actors may withhold information for ‘security reasons’. Moreover, the industry Joakim Berndtsson is studying is marked by non-transparency, with reliable information about contract volume, size or clients being hard to come by. Although these contextual gaps might be ameliorated by triangulation, this is not a guarantee. As the author suggests, when conducting interviews with representatives of individual companies, ‘security officials are often very cautious and sometimes wary of researchers and will not necessarily be open to discussions of what they feel are sensitive issues’ (Berndtsson, Chapter 6, this book).

Addressing the boundaries with respect to context, as Joakim Berndtsson but also Magnus Dau and Marlen Martin do, is useful in several respects. First, it helps to refute the myth that ‘method is and can do everything’ and that it only takes the right one or their right combination as well as adequate text sources to get at the ‘real’ story. Second, it draws attention to the relationship between methods and the research subject and how they condition each other. Most of the chapters in this section treat non-state actors as a means to an end to illustrate the method. The authors do not explicitly ask the questions that, however, seem equally relevant: what is special about non-state security actors and how may this influence or inflict upon the choice of methods? For example, is an interpretative analysis particularly suited for studying non-state actors? Does it perhaps

allow us to give voice to actors that are often voiceless and invisible, as Anja Mihr suggests? In this case, our chosen method might not simply be a tool to generate knowledge and solicit their story, but could instead also be understood as a political move. Furthermore, how do attributes or characteristics of non-state actors impede qualitative interpretative research simply because conducting interviews or engaging in participant observation is difficult or impossible to do, as it might be life-threatening to either our research subjects when exposed (e.g. members of opposition groups) or to us (e.g. in the case of terrorist groups)? Reflecting upon the fit between research objects/subjects and methods in this sense reveals that methods and their choice also involve normative and/or moral considerations in terms of the responsibility we might have vis-à-vis our research subjects.

This leads me to a last and final point related to one of the major boundaries of knowing, namely our own preconceptions. 'The assumptions with which we begin, the theories, the modes of reasoning and even our political commitments color [our] narrative and simultaneously clarify and obscure features of the past' (Froese 2013: 124). Of the authors in this section, several allude to this point. Joakim Berndtsson, for example, in reference to Christopher Kinsey (2006: 7), notes that interviewing representatives of the private security industry requires 'a certain degree of open mindedness' given the negative reputation of firms, on the one hand, and the sometimes existing scepticism on the part of researcher, on the other hand. In addition, Anja Mihr acknowledges the way in which our lenses and pre-conceived notions affect our research, which is why the 'main challenge and the task of the interviewer', according to her, '... is to control her own subjectivity and to be as neutral and reflective as possible'.

While these authors are cognizant that our epistemological lenses and identities might play a role in the research we conduct, they seem optimistic that this factor can be controlled or be kept at a minimum. As Alexander Spencer (Chapter 3, this book) notes 'through transparency, verifiability and showing the extensiveness of the narrative elements, the subjective story can be made inter-subjective'. While this certainly should be the aim, in the end and despite our care 'much, if not all, qualitative and ethnographic writing is a narrative production' (Denzin 1997: 4; see also Patterson and Monroe 1998: 324) with us telling a story of a story or as Alexander Spencer (Chapter 3, this book) acknowledges himself, 'by creating a collage of a vast number of quotes this chapter is itself a narrative'. Because of the limits imposed by our own contextuality, doing interpretation always has a double meaning. At the same time as the researcher attempts to gain insight into how political actors conceive of events or make sense of them, our epistemological lenses and identities function as filters allowing us to see or hear certain things, but not others.

Acknowledging these biases and making them explicit is what feminists have urged us to do not only because they let the reader know where we are coming from and why we may have interpreted the text in a certain way, but also these biases may be productive for interpretative analysis. As Iver Neumann points out in connection with cultural knowledge, a certain distance, however achieved,

contributes to what he refers to as a ‘marginal gaze where things look strange enough to present themselves as puzzles’ (Neumann 2008: 64). Our pre-conceived notions may precisely function in this manner, enabling us to see what is odd or different about something and needs explanation or interpretation, hence create the necessary distance.

Overall, the chapters in this section are all real gems to read since they do not simply walk us through the methods in an abstract and depersonalized sense. Instead the authors put the scholar squarely back in. All of them convey that doing research does not always follow the carefully preconceived script or plan. It hinges instead as much on the human factor – the individuals we study and ourselves – and sometimes even outright luck. For ongoing PhDs the chapters therefore should be a must read. In addition to explaining what a narrative, thematic or sentiment analysis as well as interviewing entails, the chapters introduce a healthy dose of realism into our methodological discussions. Rather than assuming that the methods we use are bullet-proof and can do magic, we should expect and be prepared for hoops and bumps along the way.

References

- Braun, Virginia and Victoria Clarke. 2006. ‘Using Thematic Analysis in Psychology’. *Qualitative Research in Psychology* 3(2): 77–101.
- Bryman, Alan. 2008. *Social Research Methods*. Oxford and New York: Oxford University Press.
- Denzin, Norman K. 1997. *Interpretive Ethnography: Ethnographic Practices for the 21st Century*. London: Sage.
- Denzin, Norman K. and Yvonna S. Lincoln. 2011. ‘Introduction: The Discipline and Practice of Qualitative Research’, pp. 1–20 in *The SAGE Handbook of Qualitative Research*, N.K. Denzin and Y.S. Lincoln, eds. Thousand Oaks, CA: SAGE.
- Froese, Marc. 2013. ‘Archival Research and Document Analysis’, pp. 118–128 in *Critical Approaches to Security: An Introduction to Theories and Methods*, L. Shepherd, ed. London and New York: Routledge.
- Hermann, Margaret G. 2008. ‘Content Analysis’, pp. 61–77 in *Qualitative Methods in International Relations: A Pluralist Guide*, A. Klotz and D. Prakash, eds. Houndmills: Palgrave Macmillan.
- Kinsey, Christopher. 2006. *Corporate Soldiers and International Security: The Rise of Private Military Companies*. London and New York: Routledge Press.
- Moon, Donald J. 1975. ‘The Logic of Political Inquiry: A Synthesis of Opposed Perspectives’, pp. 131–228 in *Handbook of Political Science, Vol. 1, Political Science: Scope and Theory* F.I. Greenstein and N.W. Polsby, eds. Reading, MA: Addison-Wesley.
- Neumann, Iver B. 2008. ‘Discourse Analysis’, pp. 61–77 in *Qualitative Methods in International Relations: A Pluralist Guide*, A. Klotz and D. Prakash, eds. Houndmills: Palgrave Macmillan.
- Pang, Bo and Lillian Lee. 2008. ‘Opinion Mining and Sentiment Analysis’. *Foundations and Trends in Information Retrieval* 2(1–2): 1–135.
- Patterson, Molly and Kristen Renwick Monroe. 1998. ‘Narrative in Political Science’. *Annual Review of Political Science* 1(1): 315–331.

Part II

Establishing causal claims



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

8 (Comparative) case studies

Combining case study techniques for the causal analysis of security privatization

Andreas Kruck

Introduction: the theory-guided study of security privatization and the usefulness of a plurality of case study techniques

Not all case studies rely on the same logic for establishing causal claims; rather, different case study methodologists (explicitly or implicitly) favour different understandings of the appropriate goals, techniques and standards for conducting case studies (see King *et al.* 1994; Van Evera 1997; George and Bennett 2005; Gerring 2007; Blatter and Blume 2008; Collier *et al.* 2010; Mahoney 2010, 2012; Blatter and Haverland 2012; Bennett and Checkel 2015). This chapter illustrates, first, how the study of non-state actors in international security can benefit from combining the case study techniques of covariational analysis, congruence analysis and causal process-tracing, and, second, how such a combination is possible in research practice. Such a combinatory approach is most useful for research that seeks to explain both broader macro trends (providing ‘the big picture’) and more specific, multicausal dynamics and policies (‘zooming in’ on the phenomenon under study). The broader purpose of this chapter is thus to demonstrate the value added in terms of reliability, breadth and depth of insights to be had from relying on all three of these case study techniques in one project. This multipronged approach is still uncommon, because covariational analysis, congruence analysis and process-tracing are usually treated as alternative approaches to case studies. In addition to proposing a triangulation of case study methods, this chapter also highlights the need to take the different modes of inference and data requirements of the three approaches seriously. Thus, such triangulation requires careful design and, most likely, additional work to gather, evaluate and weigh different types of evidence.

I illustrate the use and usefulness of a plurality of case study techniques by drawing on my own research on security privatization and private military and security companies (PMSCs). The goal of my research was to explain the use of PMSCs by Western ‘strong’ and democratic states, namely the United States, the United Kingdom, France and Germany (Kruck 2014a, 2014b). That goal led me to the combined case study approach. Although previous research offered important insights into the causes and motivations for privatizing

security, theory-guided and comparative studies have been scarce. By relying on a triangulation of case study methods not previously pursued in this research field, this project sought to help close the gap in our understanding of what drives or inhibits states to resort to the use of PMSCs.

The central question I posed, therefore, was: why have the US, the UK, France and Germany, to varying degrees, delegated security governance tasks to PMSCs? To answer this causal-analytical question, I drew on explanatory models reconstructed from extant explanations of Western security privatization (Singer 2003: 49–70; Avant 2005: 30–38; Deitelhoff 2010; Krahmann 2010: 21–83; Petersohn 2010) and on more abstract, theoretical approaches from International Relations and comparative politics, such as principal–agent theory, historical institutionalism and sociological institutionalism. Each of the three models presented a distinct causal logic behind security privatization – functional, political or ideational – from which testable hypotheses could be derived.

From a *functional* standpoint, security privatization is the rational response of legislatures and governments to changing environmental demands that challenge the problem-solving capacities and economic cost-efficiency of public (state-based) modes of security governance. This response is driven by states' resource needs and capacity gaps. Thus, functional reasoning predicts that the more complex the technological and operative contexts of warfare, the higher the dependence of public actors on PMSCs' material and/or non-material resources, and the more states will make use of PMSCs. By implication, states will first and foremost outsource highly complex tasks. Moreover, a focus on cost-efficiency leads to the functional expectation that the greater the defence budget saving pressures and the higher the economic gains from security privatization, the more states will rely on the services of PMSCs. States will first and foremost outsource tasks for which there is a competitive private market.

A *political*, power-based approach conceives of security privatization as a strategy adopted by executives to enhance their autonomy and policy influence in relation to (potential) veto players, such as national parliaments, (constitutional) courts, domestic societal actors or other governments on the international level (see Moe 1990; Wolf 1999). From this perspective, the non-transparent use of PMSCs serves to avoid politically costly scrutiny, opposition and control in the area of security policy by parliamentary bodies, civil society and the media (Singer 2003: 209–215; Avant 2005: 43, 128–133; Avant and Sigelman 2010). Thus, the riskier and the less popular a military operation is among the domestic audience, the greater the incentives of governmental actors to circumvent domestic veto players and the higher their propensity to rely on the security services provided by PMSCs. Above all, states will seek to outsource politically and societally controversial tasks to PMSCs. Moreover, the weaker the parliamentary and judicial mechanisms of oversight and control concerning governments' contracting with PMSCs, the more likely security privatization will occur and become extensive.

In an *ideational* approach, patterns of security privatization follow from the transnational diffusion of neoliberal ideas and blueprints of market-based modes

of governance, even though their implementation remains conditioned by national laissez-faire liberal or state-interventionist conceptions of the state (Krahmann 2010; Petersohn 2010; Cusumano 2014). The more the prevalent conception of the state is shaped by laissez-faire (neo-)liberal ideas and norms rather than by state-interventionist ones, the more a privatized provision of security will be accepted as legitimate, proper and even 'normal' among political decision-makers and their electorates and the more extensively the state will use the security services of PMSCs. Outsourcing of activities that affect military core functions will occur later and more reluctantly than would be the case with non-core functions, because neoliberal ideas diffuse first to the margins of armed forces' activities before they may move on into core areas of warfare.

In this project, my primary goal in analysing states' use of PMSCs was not to identify the relatively most powerful 'winning' theory (from among those three models) in a competitive theory test. First, the primary aim of my analysis was to provide a comprehensive explanation for the empirical phenomena of privatization of security. Thus, one major objective was to explain broader intertemporal and cross-country patterns, as well as more specific, country-level dynamics of the use of PMSCs. Second, the project was designed to reveal the explanatory merits and shortcomings – or the comparative advantages and weaknesses – of different theoretical models with an explicit view towards exploring rather than shutting down possibilities for theoretical synthesis.¹ For these purposes, a combination of different case study approaches proved most useful, as the remainder of this chapter will illustrate.

In the following sections, I describe the research design, focusing on the choice and application of case study methods. First, I elaborate on the selection of cases, aspects of operationalization and issues of data selection. I then introduce the different case study methods and reflect on their practical application in the project in question. I retrace how covariational analysis was employed to determine whether the functional, power-political and ideational independent variables highlighted by the alternative theories of security privatization covaried over time and across countries with the dependent variable, i.e. the use of PMSCs. This is followed by a reflection on how congruence analysis and causal-process-tracing were used for deeper and denser, theoretically saturated case studies of the US, the UK, France and Germany in order to reconstruct the more specific dynamics of security privatization. The chapter concludes with a discussion of the advantages and challenges of combining case study techniques.

Research design and methods in use: a combined case study approach

My empirical analysis relied on a structured, focused, multimethod comparison of a small number of purposefully selected cases. The analysis was structured in that I asked the same question – why has the state delegated security governance tasks to PMSCs? – of each case under study to guide and standardize data collection, thus allowing comparisons of the cases and an accumulation of the

findings; it was a focused comparison in that it addressed certain aspects and data of the cases examined, namely the delegation of security functions to PMSCs (George and Bennett 2005: 67); and it was a multimethod (case-study) approach in that it drew on covariational analysis and on congruence analysis and process-tracing.

Case selection, operationalization and data

For my case studies, I selected countries for which the national dynamics of security privatization are empirically relevant in their own right because they concern major Western states (the US, the UK, France and Germany) and (more or less) important players in international security. Another more important reason I selected these cases was that I had a theoretical and analytical goal in mind; that is, to draw conclusions about the strengths and weaknesses of different explanatory models. The selection of cases thus included data-rich cases for which not only the recent but also the historical development of privatization could be thoroughly retraced. This allowed for a longer-term perspective on privatization, thus enhancing intertemporal within-case variance. Moreover, the selection of 'diverse cases' that achieve maximum variance in terms of privatization (see Gerring 2007: 97–99) took into account that a convincing explanation should be able to explain both (very) strong security privatization tendencies (e.g. in the cases of the US and the UK) and much weaker ones (e.g. in the cases of Germany and France). Because a satisfying explanation should capture stark differences (e.g. between the US and Germany) *and* more subtle variations (e.g. between the US and the UK), the selection of cases also represented both large and small differences in privatization policies. Finally, the selection of extreme cases made it unlikely that relevant causes of privatization that might have been ignored in the three models would be overlooked, because in extreme cases such factors should be present in abundance and thus relatively easy to detect (Van Evera 1997: 80–81; Gerring 2007: 104). These selection criteria were critical to the development of a comprehensive synthetic perspective based on the reconstructed theoretical models *and* empirical evidence on security privatization in the four countries analysed.

For the operationalization and measurement of the key concept of 'security privatization', I focused on both quantitative and qualitative dimensions. This is because to understand the security privatization policies of a state and to assess how far privatization had gone, it was crucial to look at both *how many* and *what kind* of services the state bought from private contractors (e.g. waste disposal, field construction, logistics and transport, weapons maintenance, intelligence, armed protection) that were linked to military operations abroad (rather than to purely domestic security or police tasks).² The consideration of both quantitative and qualitative aspects of the phenomenon under study should be useful for a host of research projects; it also draws attention to a broad(er) range of numerical and non-numerical data sources that may serve as evidence to support one's arguments. In the study at hand, the scale of a state's reliance on PMSCs, the

scope of outsourced tasks and the depth of the state's use of PMSCs together determined the magnitude of privatization.

Apart from these general concerns about achieving a comprehensive measurement of such heterogeneous and complex phenomena as security privatization and PMSCs, the choice of more specific indicators was also influenced by my ambition to ensure comparability across cases and, quite simply, accessibility of empirical data. Thus, the ratio of (the number of) PMSC employees compared with public armed forces in zones of conflict (e.g. 1:55, 1:1, 1.5:1) and the state's budgetary expenses for private contractors (in relation to overall defence spending; e.g. 6.0 per cent, 3.0 per cent, 1.3 per cent) served as key quantitative indicators measuring the *scale* of security privatization. The *scope* and *depth* of privatization were designated by specifying the (more or less broad) range of outsourced tasks and the relative prevalence of certain types of services (e.g. logistics, weapons maintenance, armed protection). The crucial question for the qualitative depth of security privatization was whether the privatized tasks included 'merely' supporting, auxiliary and supplementary services or also essential 'core' functions for the provision of security (because the outsourcing of the latter indicates an even more fundamental and potentially more problematic transformation). Thus, a distinction was made between military core and non-core functions, with the former including combat, command and control functions; global intelligence and reconnaissance; and strategic transport and force projection (Petersohn 2006: 12–15).

Data for the case analyses were taken from policy documents and statements of decision-makers, such as governmental reports, parliamentary inquiries and reports, consultative reports by regulatory bodies, oral statements of decision-makers and the like, as well as from secondary sources (existing country studies on PMSCs) that provided a good deal of empirical information about the use of PMSCs in the US, the UK, Germany and, to a lesser extent, France. Wherever possible, I triangulated data from different primary and secondary sources to improve the reliability of my findings. Such a data double-check is particularly useful for study objects where the reliability of information from only one source (e.g. government or NGO reports) is doubtful (e.g. because the issuer of the information may have political incentives to withhold or distort relevant data).

Covariational analysis and its application: assessing the power and limits of monocausal explanations

I first used covariational analysis to observe whether functional, political and ideational explanatory variables and the dependent variable (i.e. the quantitative scale, scope and depth of security privatization through the use of PMSCs) covaried over time and across countries in order to make causal inferences and assess the explanatory strengths and weaknesses of the above functional, political and ideational propositions. An alternative approach would have been to conduct a qualitative comparative analysis (QCA) rather than a covariational analysis as a first step (see Mello, Chapter 9, in this volume). I opted for the covariational

analysis rather than the QCA because I first wanted to find out to what extent individual (functionalist, political or ideational) hypotheses might be sufficient to provide an empirically comprehensive but theoretically parsimonious explanation for the phenomenon under study. After the covariational analysis had shown that this was not the case, I subsequently engaged more thoroughly with the complexities of multicausal and dynamic explanations³ (see below). I thus sought parsimony first and then added complexity later on.

Logically similar to quantitative methods (see De Juan, Chapter 10, in this volume), covariational analysis draws causal inferences by observing whether (hypothesized) causal factors and causal effects covary, for example, over time or across countries (King *et al.* 1994; Gerring 2007). ‘If there exists such a covariance over time or space between the independent variable (X) and dependent variable (Y), we can infer that X caused Y’ (Blatter and Blume 2008: 318). The main goal of covariational analysis is to make generalizing inferences from selected cases to a wider population of cases. In this conception, relevant observations for establishing causal claims are ‘data-set observations’ (i.e. the scores for dependent and independent variables) (Gerring 2007: 20, 23–25; Collier *et al.* 2010: 184–188). If causal inferences drawn solely from covariational analysis are to be convincing, the variable-centred correlation must be interpreted by linking it to theory, and all other potentially causally relevant factors must be controlled for by choosing comparable cases and most similar cases designs (Blatter and Blume 2008: 319–320). Even under somewhat less stringent conditions, covariation of independent and dependent variables gives us some reason to believe it plausible that a theory has explanatory validity (Van Evera 1997: 58–62). In the research project described here, this approach provided an initial assessment of the explanatory plausibility of functional, political and ideational propositions. Moreover, covariational analysis is particularly useful for drawing macro pictures of broader trends and differences with regard to a certain phenomenon (e.g. security privatization), which is another reason why I applied it as a methodological ‘first cut’ in my analysis of states’ use of PMSCs.

I first turned to the dependent variable describing trends and cross-country variation in the states’ contracting with PMSCs. The idea was to provide the reader with a larger, ‘panoramic’ picture of the approaches of the US, the UK, Germany and France to contracting with PMSCs. This picture also included a sketch of the evolution of the private security industry and the broader historical context of the move from mercenary armies via the state monopoly of armed force to PMSCs. Not only did this broad history and comparative account reflect a methodological choice to prepare the ground for an assessment of causal propositions, but it was also driven by concerns about readability (avoiding the dryness of merely checking values of variables), capturing the reader’s interest in the phenomenon under study and providing a basic understanding of the phenomenon before delving deeper into more specific causal analysis and inference. But, most importantly, from a methodological perspective, I designated the scale (referring to public–private force ratios and spending for private contractors), scope (the breadth of outsourced services) and depth (the existing/non-existing

outsourcing of core military functions) of security privatization in the US, the UK, France and Germany, as well as how these values changed over time. I outlined that there are larger differences between Anglo-Saxon and continental European countries but also more subtle differences among these countries (e.g. in the range of armed services outsourced by the US and the UK).

To briefly sum up the descriptive findings, privatization is strongest among the Anglo-Saxon countries. The quantity and quality of the use of PMSCs by the US are unprecedented in modern times and are unrivalled. The UK also displays a strikingly high level of privatization in quantitative terms; at the same time, the qualitative scope of private security services is more limited in the UK. Activities outsourced by the British Ministry of Defence (MoD) focus on non-armed logistical support, which has become essential for British force projection and operability abroad. Although Germany and France have also increased their use of PMSCs, the levels of German and French security privatization are relatively low when compared with the levels in the US and UK, with Germany using a somewhat broader range of private security services than France does. The scope of privatized security services is more narrowly confined in both the continental European cases. Nonetheless, in Germany, and particularly in France, the sizable reliance on private strategic transport capacities touches on the core requirements of warfare.

I then turned to the independent variables, probing whether the causal conditions identified by the three models covary with these observable empirical trends and differences. I provided a qualitative, textual assessment of the degree of covariation between explanatory and dependent variables; that is, I explicitly reflected on which intertemporal trends and cross-national differences in the 'panoramic' picture could or could not be captured by any one of the individual hypotheses. Of course, with a larger number of cases, a similar assessment could be made by relying on quantitative, software-based techniques (see De Juan, Chapter 10, in this volume). The *functional complexity hypothesis* finds support in some limited issue areas, such as the maintenance of high-technology weaponry and strategic transport capacities where varying levels of dependencies on PMSCs have emerged, which underlie variations in states' use of private security services. Despite these insights, the complexity hypothesis also faces serious problems. Although the impression that recent asymmetric conflicts and new military technology have rendered warfare more complex is quite common among policymakers and academic observers (Ballard 2007), there is little factual evidence in intertemporal and cross-country data that the operational and technological complexity of warfare is *the major driver* of broader privatization trends. First of all, one may reasonably doubt that contemporary violent conflicts are indeed operationally and technologically more complex than previous security threats were (e.g. in the context of the confrontation of nuclear superpowers during the Cold War) (Biddle 2004; Mello 2010). More specifically, the statistics on states' defence budgets provided by NATO and the Stockholm International Peace Research Institute (SIPRI) indicate that the relative share of expenses for armament technology in the defence budgets of the analysed

countries has remained fairly constant for the past 30 years. At any rate, there is no robust evidence that the percentage of spending on complex high-technology goods and activities relative to other defence expenditures has increased in a way that is commensurate with the increased use of PMSCs. Finally, a large proportion of PMSCs do not perform complex high-tech tasks such as the maintenance and operation of sophisticated means of warfare. Rather, they offer simple, low-skill services, such as cleaning, construction work, waste disposal and other day-to-day camp management tasks. In those large areas of low-skill services, resource dependencies do not play a major role.

As for the functional *cost-efficiency hypothesis*, NATO and SIPRI data show that the defence budgets of the US, the UK, Germany and France were indeed all reduced after the Cold War. However, within the past ten years, defence spending has soared again (especially in the US) while the use of PMSCs has simultaneously grown. This finding is at odds with the expectation that defence budget pressures (i.e. cuts of defence budgets in relation to gross domestic products) drive privatization. Moreover, defence budget pressures have been stronger in Germany than in the US since the Cold War, but, in contrast to the US, Germany has not pursued a comprehensive, supposedly cost-saving strategy of security privatization. The economic cost-efficiency hypothesis also has a hard time accommodating findings that suggest that claims of reducing costs and increasing cost-efficiency through outsourcing have frequently failed to materialize (Deitelhoff 2010: 194–195; Petersohn 2010: 545–546). As of 2005, about 40 per cent of US Department of Defense (DoD) contracts were awarded on a non-competitive basis, often with ‘cost plus’ arrangements, which guarantee companies more profit if they spend more state money in the fulfilment of their contractual tasks. This hardly speaks for a (growing) preference of states to outsource to competitive private markets.

The *political* model appears to be supported by numerous studies that found indications that the use of PMSCs is indeed harmful to the transparency, accountability and parliamentary control of military operations and that governments are at the very least aware of this (Avant 2005; Avant and Sigelmann 2010; Deitelhoff 2010). However, even a cursory glance at poll data on public support for major military operations (from the Korean to the Afghanistan War) in which the US, the UK, France and/or Germany were involved, reveals that the (un-)popularity of military operations and the use of private contractors do not covary systematically over time and across countries. In more general terms, it is not evident that incentives for Western states to reduce political costs should have substantially increased over the past two decades while the use of PMSCs has soared. The political model also leads to erroneous cross-national expectations. One would expect that states such as Germany, which have had less experience with fallen soldiers in the recent past and are thus more ‘casualty-sensitive’, rely more heavily on (armed) PMSCs than do states such as the US, which have regularly lost a substantial number of soldiers’ lives in combat. This assumption flatly contradicts empirical patterns of security privatization (Petersohn 2010: 545). Moreover, if political costs as such were the main determinant

of security privatization, it should be puzzling that hardly controversial tasks such as logistics and maintenance work have been privatized extensively, whereas politically costly casualty-prone tasks such as combat-related activities have hardly been privatized at all. In other words, the model is of little help in explaining the delegation of uncontroversial logistical tasks to PMSCs, which is one of the most prevalent areas of privatization.

The *ideational model* is supported in that conceptions of the state and the quantity and quality of states' use of PMSCs correlate strongly. In the past two decades, neoliberal policy ideas and market-based modes of governance have been on the rise across numerous issue areas in the member states of the Organisation for Economic Cooperation and Development (OECD). The transnational diffusion of a neoliberal model of lean armed forces is reflected in the common privatization strategies that have been included in the security and defence strategies of nearly all Western OECD states. Nonetheless, the prevailing (national) laissez-faire liberal or state-interventionist conceptions of the state, and consequently ideas-based predispositions to outsource core policy functions, such as the provision of security, still differ considerably: the US and the UK are characterized by a more laissez-faire liberal and less state-interventionist conception of the state than France and Germany (Prasad 2006). This finding is corroborated by indices of economic freedom such as the ones provided by the Canadian Economic Freedom Network and the Heritage Foundation in the US (Gwartney *et al.* 2011 and previous annual reports; Heritage Foundation 2012; see also Petersohn 2010: 540–541 for a similar argument). In fact, overall government size indicators from these two indices for the US, the UK, Germany and France largely *covary* (over time and across countries) with the scale and scope of the respective states' use of PMSCs.

Under the privatization-friendly macro-ideational conditions in the UK, and especially in the US, the use of PMSCs is perceived as more acceptable and 'normal', and closely tied public–private privatization coalitions have thrived. The large-scale use of PMSCs in the US and the UK can be perceived as the emanation and continuation of broader neoliberal thinking about defence and the changing trends in defence economics that increasingly took hold in both countries in the early and mid-1990s. Reliance on PMSCs in the US and the UK is not an isolated phenomenon but has followed from the increasing influence of neoliberal ideas on defence economics and appropriate modes of security provision, which never took root in a comparable way in continental European states. This evolution is also underlined by existing studies tracing the differential impact of ideologies, norms and ideas on security privatization in the US, the UK and Germany (Krahmann 2010; Petersohn 2010). However, the ideational model also leaves some blind spots. It has little to say about the exact *timing* and *speed* of the diffusion of neoliberal privatization ideas into the realm of security. Although neoliberalism has shaped political culture in the UK and the US at least since the Thatcher and Reagan administrations in the early 1980s, it was only in the mid-1990s, under the Major and Clinton administrations, that the two countries extensively and systematically outsourced security services to PMSCs

at home and abroad. Moreover, from a purely ideational perspective, the differences in the quantity and particularly the quality of privatization between the US and the UK should be even smaller than we observe today. The different policies of the US DoD and the British MoD concerning the use of armed PMSCs can hardly be explained by different laissez-faire liberal or state-interventionist conceptions of the state. Moreover, Germany's and especially France's notable use of private capacities in specific issue areas (e.g. air transport) cannot be captured by ideational reasoning.

So, after all, the covariational assessment suggested that none of the models had comprehensive and sufficient explanatory power on their own. But, despite their blind spots, all of these models provided conditions and mechanisms that appear, more or less, relevant for a comprehensive explanation. More specifically, the analysis revealed that all three of the models had particular explanatory strengths and weaknesses. Zooming in from this broader comparative picture to the more specific country-level dynamics of security privatization, I went on to draw on causal-process-tracing and congruence analysis in four country case studies of the outsourcing of security functions to PMSCs by the US, the UK, France and Germany. These techniques were chosen because, unlike covariational analysis or other medium- or large-*n* approaches, they provide deeper and denser insights into the dynamic unfolding and interaction of causal mechanisms over time.

Congruence analysis and process-tracing and their application: towards a more dynamic and synthetic explanation

Process-tracing focuses on retracing causal mechanisms; that is, the (inter-)linkages and the unfolding of causal factors over time that produce a certain case outcome. With its case-centred approach, process-tracing looks at 'causal process observations' (Collier *et al.* 2010; see also George and Bennett 2005; Mahoney 2010), 'the observable within-case implications of causal mechanisms' (George and Bennett 2005: 138) and, potentially, the complex interaction between different causal factors in the process (see also Mello, Chapter 9, in this volume). This involves not only checking independent and dependent variables but also searching for 'traces for every step between the cause and the outcome' (Blatter and Blume 2008: 320). We have to 'trace [...] backward the causal process that produces the case outcome, at each stage inferring from the context what caused each cause. If this backward process trace succeeds, it leads the investigator back to a prime cause' (Van Evera 1997: 70). In other words, this is about uncovering the chain of events or the decision-making process by which initial conditions are translated into case outcomes (George and Bennett 2005: 217). Thus, process-tracing analysts rely on the reconstruction of dense and detailed historical narratives (Büthe 2002).⁴

Congruence analysis draws inferences from the congruence or non-congruence of concrete observations with specified predictions from abstract theories to the relevance of these theories for explaining the case or cases under

study (Blatter and Blume 2008: 325; see also Van Evera 1997: 56 and George and Bennett 2005 for elements of a congruence analysis approach). In contrast to covariational analysis, congruence analysis uses a much broader set of predictions and observations beyond the values of independent and dependent variables. This includes not only causal process observations (e.g. sequences of actions and events) but also expectations about the most important actors, their perceptions and their motivations, structural factors and so on. Thus, congruence analysis requires rich, complex (rather than parsimonious) and clearly framed theories that contain a large number of observable implications; in this conception, a large number of causally relevant factors specified by the theory is an asset rather than a liability for making causal inferences. The more matches between theoretical prediction (model components) and empirical observation (empirical case study elements), the greater the explanatory power of the model (Blatter and Blume 2008: 326–328; Blatter and Haverland 2012: 144–204). However, the attribution of concrete observations to abstract concepts and to theoretical frameworks is not simply a technical manner (of checking fully operationalized variables) but requires explicit and extensive reflection, interpretation and justification. Interpretative techniques are necessary to make the ‘inferential leap’ from specific observations to abstract theories.

Congruence analysis thus contributed to a denser and theoretically more saturated account of security privatization situated in a broader theoretical discourse about the privatization of governance. Causal-process-tracing in turn helped me to trace deeper causal mechanisms of security privatization over time, because it focused on the temporal unfolding and interaction of different functional, political and ideational factors. The combination of congruence analysis and causal-process-tracing allowed me to assess the more specific explanatory strengths and weaknesses of the different models and to retrace the institutional dynamics of security privatization, as well as the causal workings of functional, political and ideational mechanisms in the US, the UK, France and Germany, thus drawing attention to important aspects that were hardly captured by mere covariational analysis. Based on these country case studies, I was able to make more fine-grained causal inferences about security privatization, point out insights and blind spots of the three explanations with regard to the process of security privatization, analyse the interaction of different explanatory variables and thus draw a number of lessons for theoretical debates about delegation in security governance.

The basic procedure for each of the country case studies was to produce a dense and detailed historical narrative of several thousand words of the process of security privatization. For example, in the case of the US, I retraced how ideational legacies from the eighteenth-century revolutionary wars shaped post-Second World War decisions in the 1950s and 1960s regarding public–private competition, which in turn paved the way for privatization policies under the Reagan administration. These decisions were then followed by post-Cold War congressional and executive initiatives to broaden the scope of security privatization. The process culminated in the George W. Bush administration’s defence

strategy, which called for the privatization of all activities not directly linked to fighting wars and led to the unprecedented use of PMSCs in the Iraq and Afghanistan wars. This thick reconstruction, produced for each of the four country cases, was not merely descriptive, it also provided a causal interpretation of how different explanatory factors shaped political decision-making at critical stages of the historical process; that is, at all major junctures where the course for or against (further) security privatization was set.

Thus, I retraced the move towards the use of PMSCs, including its drivers and brakes, for all four countries. I then undertook an explicit comparative interpretation of the match between those case narratives and the propositions of the functional, political and ideational models. I should stress here that whereas the covariational analysis followed a more standardized logic of measuring values of variables (previously operationalized with the help of specific, a priori defined indicators), the process-tracing and congruence analyses complemented this covariational search for broad causal trends and patterns with a deeper and more interpretive approach that is better suited to uncovering complex and more specific dynamics of a given phenomenon. In essence, I came up with a causal *interpretation* of the process and practices of security privatization, the plausibility of which the interested reader should be able to evaluate on the basis of empirically rich and extensive, if not 'complete', country case studies.

What kind of theoretical inferences could be gained from the process-tracing and congruence analyses? Similar to the covariational analysis, these analyses revealed a mixed record of the three distinct models, which captured some aspects of security privatization but failed to grasp others. More importantly for comprehensively explaining the phenomenon under study and arguably also for theoretical progress in the analysis of the privatization of governance functions, the case studies strongly suggested that functional, political and ideational conditions and mechanisms of security privatization did not operate independently of one another but rather reinforced or blocked one another's operation. The country profiles lent considerable support to the key argument that functional triggers, in the form of states' capability gaps and resulting dependencies on PMSCs, are important triggers of contractual institutional change, but their operation will be heavily conditioned by political and ideational variables. Large-scale and broad-ranging use of PMSCs is likely when (1) privatization-friendly ideas are already in place in the country under study; when (2) (strong) functional triggers create incentives for security privatization, with states looking for domestic historical precedents and/or contemporary exemplars in other countries in order to choose how to deal with these functional challenges; and when (3) political brakes on the use of PMSCs are absent or inactive.

This combination of conditions and mechanisms – which, framed in abstract models, could be applied to and tested against other cases of privatization – captures the dynamics and differences in and between the country cases very well. In the US, functional triggers for increased contracting after the Cold War met with well-established neoliberal idea systems that favoured a strong role of the market in the provision of collective goods, whereas political brakes in the

form of constitutional limits or political contestation of security privatization were absent. Moreover, US history has been rich in precedents indicating a reliance on private forces and private support to armed forces. As a result, the US went in for the extensive use of PMSCs from the mid-1990s. In the UK, similar functional and ideational conditions for extensive security privatization have been in place, but political costs and contestation of politically delicate relationships between the government and PMSCs have constrained the use of PMSCs, at least in some controversial issue areas such as armed security and foreign military training. The Cold War mercenary roots of many British firms, as well as scandals in the 1990s (most notably the Arms to Africa Affair) have put a brake on some privatization drives. France and Germany have also faced functional incentives to rely on PMSCs to meet new security challenges and cut down the size of their armies. But France and Germany have only hesitantly and reluctantly followed the Anglo-Saxon model because of their ideational legacies of state-interventionist conceptions of the state and a strict state monopoly on the use of force. Activated political brakes in the form of constitutional limits and political resistance to a high-profile military role for the country (in the German case) and privatization of sovereign functions (in both cases) also play an important role. A model of extensive security privatization would thus be a bad fit with their ideational traditions and would be politically costly. The German government's reluctance to publicly embrace the needs-based contracting that has occurred, as well as the wait-and-see approach to regulation, testifies to this fundamental uneasiness. France's more overt ambitions to play an active structuring role in the security industry through co-ownership and nurturing of national champions is in line with French conceptions of an active role for the state in economic governance.

Conclusion: value added and the challenges of combining case study techniques

The combination of covariational, process-tracing and congruence-analysis techniques yielded more, i.e. broader, deeper, more differentiated and complementary, insights that I could not have gained from just a covariational or process-tracing analysis. Embracing the plurality of existing case study techniques and combining their strengths promises to enhance the range, richness and reliability of causal inferences that can be drawn from (comparative) case studies of non-state actors in international security.

But, of course, such a combination also poses challenges. It requires quite a bit of work to collect, evaluate and interpret many different kinds of empirical data and points of observation. More fundamentally, combining these techniques should not imply that differences in their logics are ignored (Blatter and Blume 2008; Collier *et al.* 2010; Mahoney 2010, 2012). Covariational analysis is essentially wedded to a quantitative understanding of drawing causal inferences from the covariation of independent and dependent variables (i.e. a very limited range of dataset observations) over as many points of observation (cases) as possible,

while ideally keeping all other factors constant. By contrast, the logics of congruence-analysis and causal-process tracing require deep involvement with and detailed knowledge of a much broader set of causal-process observations (beyond the independent and dependent variables) for a limited number of well-selected cases (e.g. 'hard cases' for confirming a theory and 'easy cases' for refuting a theory when the goal of research is a theory test). Here, causal leverage does not come from increasing the number of cases but from taking a very close look at the (very) few chosen cases. These different logics imply different data needs and divergent understandings of the role of operationalization and interpretation for making causal inferences. Whereas covariational analysis typically checks the values for *a priori* operationalized variables (i.e. their indicators) at several points of observation, congruence analysis and causal process-tracing involve (explicit) causal interpretation; that is, reflection on what a certain causal-process observation means for the explanatory power of particular theories.

These different logics could be taken into account during the distinct methodological steps of the described research project on states' use of PMSCs (which relied *first* on covariational analysis, *then* on process-tracing and *ultimately* on congruence analysis). But for the sake of coherence of the overall study and to arrive at common, additive inferences (such as the above synthetic explanatory argument) through different methodological strategies and via heterogeneous empirical data, it was necessary at times to make tough choices, privileging one kind of evidence over another (e.g. evidence from the historical narratives over cross-country covariation) and compromises that led to the application of less than 'pure' and 'ideal' (covariational, congruence-analysis and process-tracing) case study logics. This is also evident in the number and selection of cases and in the number of causal-process observations that could be covered.

Such compromises might be a cause for concern for the purist methodologist wedded to a certain understanding of how to carry out (comparative) case studies. But, from the perspective of problem-oriented research interested in a comprehensive understanding of a complex and multifaceted phenomenon, a combination of various approaches, not only on the theoretical level but also on the methodological level, appears to be a fruitful strategy for studying non-state actors in international security.

Notes

- 1 Even if monocausal theories might ultimately be considered lacking when the analytical objective is to comprehensively grasp complex, multifaceted phenomena such as security privatization, a very reasonable way to achieve a systematic theoretical synthesis still departs from the deductive reconstruction of distinct, internally consistent models and their application to empirical cases (see Legro and Moravcsik 1999: 50).
- 2 Although this conceptual distinction between domestic/security and international/defence-related contracting (with my research focusing on the latter) is more blurry in the empirical world of security privatization, it was important to keep this distinction in

mind when comparing data from different sources and countries that might vary in their understanding of what qualified as security privatization (including or excluding domestic privatization of police functions in Western OECD states).

- 3 Moreover, QCA is good for studying multicausal relationships but is less well suited to researching dynamic causality, which turned out to be quite important in my study.
- 4 For a proposal and a detailed elaboration of process-tracing as a rigorous scientific method for uncovering necessary causal connections, see Bennett and Checkel (2015); for a distinction of different kinds of process tracing tests for causal inference, such as straw-in-the-wind, hoop and smoking-gun tests, see Collier (2011).

References

- Avant, Deborah. 2005. *The Market for Force: The Consequences of Privatizing Security*. Cambridge, UK: Cambridge University Press.
- Avant, Deborah, and Lee Sigelmann. 2010. 'Private Security and Democracy: Lessons from the US in Iraq'. *Security Studies* 19(2): 230–265.
- Ballard, Kyle M. 2007. 'The Privatization of Military Affairs: A Historical Look into the Evolution of the Private Military Industry', pp. 37–53 in *Private Military and Security Companies: Chances, Problems, Pitfalls and Prospects*, T. Jäger and G. Kümmel, eds. Wiesbaden: Verlag für Sozialwissenschaften.
- Bennett, Andrew, and Jeffrey T. Checkel, eds. 2015. *Process Tracing: From Metaphor to Analytic Tool*. Cambridge, UK: Cambridge University Press.
- Biddle, Stephen D. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton: Princeton University Press.
- Blatter, Joachim, and Till Blume. 2008. 'In Search of Co-Variance, Causal Mechanisms or Congruence? Towards a Plural Understanding of Case Studies'. *Swiss Political Science Review* 14(2): 315–356.
- Blatter, Joachim, and Markus Haverland. 2012. *Designing Case Studies: Explanatory Approaches in Small-N Research*. Houndmills: Palgrave Macmillan.
- Büthe, Tim. 2002. 'Taking Temporality Seriously: Modeling History and the Use of Narratives as Evidence'. *American Political Science Review* 96(3): 481–493.
- Collier, David. 2011. 'Understanding Process Tracing'. *Political Science and Politics* 44(4): 823–830.
- Collier, David, Henry E. Brady, and Jason Seawright. 2010. 'Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology', pp. 161–199 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edn. H.E. Brady and D. Collier, eds. Lanham: Rowman & Littlefield.
- Cusumano, Eugenio. 2014. 'The Scope of Military Privatisation: Military Role Conceptions and Contractor Support in the United States and the United Kingdom'. *International Relations*. doi:10.1177/0047117814552142.
- Deitelhoff, Nicole. 2010. 'Private Security and Military Companies: The Other Side of Business and Conflict', pp. 177–201 in *Corporate Security Responsibility? Corporate Governance Contributions to Peace and Security in Zones of Conflict*, N. Deitelhoff and K.D. Wolf, eds. Houndmills: Palgrave Macmillan.
- George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Gerring, John. 2007. *Case Study Research: Principles and Practices*. Cambridge, UK: Cambridge University Press.
- Gwartney, James, Robert Lawson, and Joshua Hall. 2011. *Economic Freedom of the World: 2011 Annual Report*. Vancouver: Fraser Institute.

- Heritage Foundation. 2012. 'Country Rankings'. Available from: www.heritage.org/index/ranking (accessed 1 March 2015).
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Krahmann, Elke. 2010. *States, Citizens and the Privatization of Security*. Cambridge, UK: Cambridge University Press.
- Kruck, Andreas. 2014a. *States and Private Authority in Security and Finance: Delegation, Regulation and the Struggle for Control*. PhD dissertation, Munich: Ludwig Maximilian University of Munich.
- Kruck, Andreas. 2014b. 'Theorising the Use of Private Military and Security Companies: A Synthetic Perspective'. *Journal of International Relations and Development* 17: 112–141.
- Legro, Jeffrey W., and Andrew Moravcsik. 1999. 'Is Anybody Still a Realist?' *International Security* 24(2): 5–55.
- Mahoney, James. 2010. 'After KKV: The New Methodology of Qualitative Research'. *World Politics* 62(1): 120–147.
- Mahoney, James. 2012. 'The Logic of Process Tracing Tests in the Social Sciences'. *Sociological Methods & Research*, 41(4): 570–597.
- Mello, Patrick. 2010. 'In Search of New Wars: The Debate about a Transformation of War'. *European Journal of International Relations*, 16(2): 297–309.
- Moe, Terry M. 1990. 'Political Institutions: The Neglected Side of the Story'. *Journal of Law, Economics, and Organization* 6(S1): 213–253.
- Petersohn, Ulrich. 2006. *Die Nutzung privater Militärfirmen durch US-Streitkräfte und Bundeswehr*. Berlin: German Institute for International and Security Affairs.
- Petersohn, Ulrich. 2010. 'Sovereignty and Privatizing the Military: An Institutional Explanation'. *Contemporary Security Policy* 31(3): 531–552.
- Prasad, Monica. 2006. *The Politics of Free Markets: The Rise of Neoliberal Economic Policies in Britain, France, Germany, & the United States*. Chicago: Chicago University Press.
- Singer, Peter W. 2003. *Corporate Warriors: The Rise of the Privatized Military Industry*. Ithaca: Cornell University Press.
- Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca: Cornell University Press.
- Wolf, Klaus D. 1999. 'The New Raison d'État as a Problem for Democracy in World Society'. *European Journal of International Relations* 5(3): 333–363.

9 Qualitative comparative analysis and the study of non-state actors

Patrick A. Mello

Introduction

Suppose you wanted to find out under what conditions non-state actors (NSAs) rebel against state authorities or to investigate what factors lead governments to contract private military and security companies (PMSCs). For these and similar research aims there will rarely be a single cause to account for the outcome. Instead, you might discover that multiple factors bring about the phenomenon of interest and that these interact in specific ways. Qualitative Comparative Analysis (QCA) is ideally suited to analyse these kinds of causal relations, especially if the aim is to conduct a comparative study of at least a medium number of cases. Two particular strengths of QCA are its ability to account for *equifinality* and *conjunctural causation*. The first concept relates to the potential presence of alternate pathways towards an outcome, and the second concerns the idea that configurations of conditions can be jointly necessary and/or sufficient for an outcome. *Fuzzy sets* complement QCA as a methodological tool for translating categorical concepts into measurable conditions, drawing on the notion that cases can hold degrees of membership in a given set.

QCA was introduced and further developed by the sociologist Charles Ragin (1987, 2000, 2008). Recent textbooks indicate that QCA has gained recognition among social scientists as a methodological approach that offers specific benefits for comparative studies (Blatter and Haverland 2012; Gerring 2012; Rohlfing 2012; Schneider and Wagemann 2012). From its inception, QCA was aimed at the ‘middle ground’ between quantitative and qualitative methodologies (Ragin 2000: 22). In contrast to regression-based statistical approaches, QCA is based on *set theory*; as such, it investigates the specific conditions under which an outcome occurs rather than estimating the ‘average effects of independent variables’ (Mahoney 2010: 132). Hence, causal relations are expressed in terms of necessary and sufficient conditions – a ‘substantively important’ view of causation that has recently gained increased attention in the social sciences (Collier *et al.* 2010: 147).

Most of the existing QCA applications in international security, and in International Relations (IR) more generally, compare countries or national governments as their cases or units of observation. Hence, QCA has at times been

equated with an approach that is exclusively macro-comparative. Contrary to this view, however, there are no inherent methodological reasons why QCA cannot be applied to the study of NSAs. Equifinality and conjunctural causation are just as relevant when it comes to researching NSAs in international security. For example, many of the explanations for the outbreak or termination of civil war or states' use of PMSCs combine several conditions and specify alternative pathways to the same outcome. Moreover, as a rigorous comparative method, QCA offers advantages over some alternative approaches.

Yet, as outlined by Andreas Kruck and Andrea Schneider in the introduction to this volume (see Chapter 1), there are challenges when it comes to research on NSAs in international security, specifically when the aim is to conduct a *comparative* study. Without doubt, the greatest of these is the problem of finding *reliable* and *comparable* data on NSAs. This problem also exists in other substantive fields, but the security realm is notorious for having inaccessible and unreliable data, which is why many studies in this area have focused on a single case or a small number of cases, based on in-depth qualitative information. However, an advantage of QCA is that researchers can draw equally on qualitative *and* quantitative sources of data, depending on their specific research aim and the kind of information required to operationalize their theoretical framework. Whereas statistical approaches require standardized data, QCA allows for more flexibility on the part of the researcher in terms of using unstandardized sources of information or case-specific indicators. This can be achieved through the calibration of crisp and fuzzy sets, a process that forms the backbone of any set-theoretic analysis. Thus, although QCA has been rarely used to study NSAs in the field of international security, it holds considerable potential as a way to enrich the portfolio of methods currently employed in this research area.

The aim of this chapter is twofold. The first part introduces the methodological approach of QCA, including its core principles, specific terminology and analytical procedures. While this methodological discussion may appear rather dense to readers unfamiliar with this approach, the second part of the chapter is intended to illuminate the research process of a study that used QCA, showing step-by-step how some of the somewhat abstract 'nuts and bolts' translate into an empirical study. This second part draws on my own work on democratic participation in armed conflict (Mello 2012, 2014). The demonstration of QCA in use should provide helpful guidance for conducting medium-*N* research on NSAs. The final section briefly covers some of the advantages and challenges of applying this analytical method.

Principles and terminology of QCA

Crisp and fuzzy sets

Although there are several variants of QCA, the most popular ones are crisp-set QCA and, increasingly so, fuzzy-set QCA (fsQCA). Crisp sets take on binary values, whereas fuzzy sets can take on any value between 0 and 1. Which of

these should be applied will depend on the research aim of a given project, but in many cases there will be advantages to using fuzzy sets. Fuzzy-set theory was developed by Lotfi Zadeh (1965) as an extension of traditional set theory. Based on the notion that cases can hold degrees of membership in a given set, fuzzy sets enable the researcher to translate categorical concepts into measurable conditions. Thus, fsQCA allows for qualitative differentiation: on the basis of substantive and theoretical knowledge, the researcher establishes ‘qualitative anchors’ to determine whether a case is ‘fully in’ a given set (fuzzy score 1), whether it is ‘neither in nor out’ (fuzzy score 0.5) and at what point a case is ‘fully out’ of a set (fuzzy score 0). This set-theoretic conception and calibration procedure challenges an assumption often made in statistical research; that is, where all variation is held to be equally meaningful (Ragin 2000: 163).

Three different procedures are possible for the *calibration* of fuzzy sets. In the straightforward approach, fuzzy scores are assigned to cases on the basis of substantive and theoretical knowledge. Using this approach, a researcher would first conceptualize different degrees of membership in a given set and then qualitatively assess the fuzzy score of each case. Other coding procedures are the ‘direct’ and ‘indirect’ methods of calibration (Ragin 2008: 85–105), which become relevant only when quantitative data are used. For example, a study interested in the severity of non-state armed conflicts could draw on various forms of qualitative information (interviews, reports, secondary sources, etc.) or could use an existing data set, such as the information provided by the Uppsala Conflict Data Program (UCDP), which could prove particularly beneficial when the aim is to conduct a larger comparative study.

Both direct and indirect methods use statistical estimation techniques to transform interval-scale variables into fuzzy-set scores. The direct method of calibration applies a logistic function to transform raw data into fuzzy-set values based on three qualitative breakpoints specified by the researcher (Ragin 2008: 89–94). As the name implies, the indirect method of calibration includes an additional step that necessitates a preliminary qualitative grouping of cases according to their degree of membership. In turn, a fractional logit model is used to estimate fuzzy-set values based on the raw data and the initial qualitative coding of cases (Ragin 2008: 94–97). It is apparent from these procedures that, despite their dissimilarities, all calibration procedures require careful *conceptualization* of qualitative anchors. Hence, even for ‘semi-automated calibration techniques’, substantive knowledge is a prerequisite for the coding of fuzzy sets (Schneider and Wagemann 2012: 41).¹

Boolean algebra

QCA is grounded in Boolean algebra, originated by George Boole, a nineteenth-century British mathematician and logician who developed an algebra for variables that have only two possible values: true (present) or false (absent). By convention, QCA solution terms are expressed in Boolean notation, which comprises several basic operators. Explanatory and outcome conditions are stated in

capital letters, whereas a tilde [\sim] refers to a logical NOT, as in the negation or absence of a condition. Multiplication [$*$] refers to a logical AND, or the combination of conditions, whereas addition [$+$] indicates a logical OR, as in alternative pathways. Finally, arrows express the relationship between one or several explanatory conditions and the outcome. Accordingly, a rightwards arrow [\rightarrow] refers to a sufficient condition, whereas a leftwards arrow [\leftarrow] indicates a necessary condition.

For example, the civil war literature considers economic motives [E] and injustice [I] to be drivers of rebellion. However, to be sufficient for the outbreak of civil war [W], these factors must combine with a potential rebel group's opportunity to form an armed movement, which depends on the presence of financial resources [R] and popular support [S]. In Boolean notation, these alternative pathways would be expressed as: $(E*R*S)+(I*R*S)\rightarrow W$. This indicates that the conjunctions ERS and IRS are both sufficient for the outbreak of civil war.

Set-theoretic methods are governed by three simple mathematical principles that can be applied equally to crisp sets and fuzzy sets.

First, the *negation* of set values is calculated by subtracting the membership value of a case in a given set from 1. If case A holds a fuzzy membership value of 0.3 in set Y, then its value for $\sim Y$ is 0.7.

Second, the combination of conditions, a *logical AND*, refers to the minimum membership values in the respective conditions. For instance, assuming we wanted to calculate the membership of case A in the combination of the fuzzy sets B and C. If A's membership in B were 1.0 (fully in the set) but its set membership in C were 0.2 (mostly outside the set), then membership in the combination of these conditions $B*C$ would be the lower of the two values (0.2). This set-theoretic principle contrasts with quantitative approaches that would calculate the *average* value between the two conditions. However, with regard to their membership in a given combination, no difference exists between cases with membership scores outside only one or both sets of the respective combination.

Finally, the third principle is the *logical OR*, which reflects the presence of alternative conditions and thus refers to the *maximum* of the respective membership values. For example, this could be the case when two conditions, A and B, individually lead towards an outcome. Hence, a case's fuzzy-set membership in the term $(A+B)$, as in 'A OR B', refers to the case's highest membership score across the two conditions.

Complex causation

As a set-theoretic method, QCA interprets relationships between social phenomena in terms of necessary and sufficient conditions. This perspective on causal relations entails three methodological assumptions: equifinality, conjunctural causation and causal asymmetry. Together, these constitute 'complex causation', a specific asset of QCA (Ragin 2008: 78; Schneider and Wagemann 2012). In a nutshell, QCA allows for the possibility that different pathways

Table 9.1 INUS and SUIN causes

<i>A and B as INUS causes</i>	<i>F and G as SUIN causes</i>
$AB + C \rightarrow Y_1$	$D \leftarrow Y_2$ $F + G \rightarrow D$

Source: own table.

towards the same outcome exist (*equifinality*); that two or more conditions can jointly cause an outcome to occur (*conjunctural causation*); and that an identified relationship between a condition and the outcome does not mean that the inverse relationship must also be true (*causal asymmetry*).

In social science research, however, it is apparent that conditions are seldom *individually* necessary and/or sufficient for an outcome. By contrast, ‘INUS’ and ‘SUIN’ causes are found frequently, although causal explanations are not often framed in these terms. INUS denotes ‘an *insufficient* but *necessary* part of a condition, which is itself *unnecessary* but *sufficient* for the result’ (Mackie 1965: 245, emphases in the original). SUIN refers to ‘a *sufficient* but *unnecessary* part of a factor that is *insufficient* but *necessary* for an outcome’ (Mahoney *et al.* 2009, emphases in the original). It follows from the definition of INUS causes that these are present whenever equifinality and conjunctural causation *combine*, meaning that at least two pathways towards an outcome exist and that at least one of them comprises more than a single condition. In turn, SUIN causes can be understood as equivalent indicators that *constitute* a necessary condition, where each individual element is unnecessary but sufficient for the condition. Table 9.1 shows the relationship in formal notation. The conditions A and B are INUS causes for the outcome Y_1 ; the conditions F and G are sufficient conditions for D, which is a necessary condition for the outcome Y_2 .

The third methodological assumption, causal asymmetry, implies that the solution for the *non*-outcome cannot with certainty be derived from the solution for the outcome. Therefore, it is considered ‘good practice’ to conduct separate analyses for the outcome and its negation (Schneider and Wagemann 2010: 408–409). This procedure can also serve to validate a theoretical claim – if a specific conjunction leads consistently towards the outcome but *also* leads towards the non-outcome, doubts arise about its explanatory strength. Moreover, a meaningful analysis of the non-outcome requires the inclusion of *negative cases*, which can strengthen confidence in the QCA results for both analyses (Mello 2013: 13–14).

Measures of consistency and coverage

Fuzzy-set analysis introduces the measures of *consistency* and *coverage* to assess whether a single condition or a conjunction of several conditions is necessary and/or sufficient for an outcome.² Whereas consistency reflects the *fit* of the

empirical evidence with an assumed set-theoretic relationship, coverage indicates the *relevancy* of a condition in empirical terms. For the analysis of sufficiency, set-theoretic *consistency* indicates the extent to which instances of a combination of conditions are a subset of instances of the outcome. Formally speaking, if all values for Y are equal to or less than their corresponding values for X, then Y is a subset of X and thus X is a *necessary condition* for the outcome. In turn, if all values for X are equal to or less than their corresponding values for Y, then X resembles a subset of Y and is thus a *sufficient condition* for that outcome.

The calculation of set-theoretic *coverage* is inversely related to the consistency measure. This implies that the coverage of a sufficient combination of conditions indicates the size of the empirical overlap, or the proportion of instances of the outcome that are explained by the given combination. Although conjunctions with several conditions are likely to show higher consistency scores, their empirical relevance will tend to decrease, because there will be fewer empirical cases that fit the described causal path. In turn, for a necessary condition, the coverage value reflects the fit between instances of that condition and the outcome. Even though a condition could be a perfectly consistent superset and thus a necessary condition in formal terms, it may be *irrelevant* in theoretical terms if the condition is present across cases that show the outcome and among cases that do not show the outcome.

Truth table analysis

How do these concepts and principles translate into the analysis? The core of the QCA procedure contains two steps. First, a *truth table* is constructed that contains rows for each logical combination of conditions and indicates which cases belong to a configuration and how these relate to the outcome. Hence, the fuzzy-set truth table represents a multidimensional vector space with 2^k corners, where k relates to the number of conditions and each corner of the resulting property space signifies a distinct combination of conditions, represented by a separate row in the truth table. For example, if four conditions are part of the analysis, the truth table comprises 2^4 , or 16, rows. Based on their fuzzy-set membership scores for each condition in a respective combination, cases can be assigned to distinct corners of the property space (Ragin 2008: 124–135). The consistency column of the truth table indicates the extent to which the fuzzy-set values of all cases in a given row or conjunction are sufficient for the outcome. Based on the consistency scores, the researcher determines a cut-off point to indicate which rows will be included for the remainder of the truth table procedure (Ragin 2008: 135).

The second step involves the *logical minimization* of the truth table, which is required to identify sufficient conditions. Here, Boolean algebra is applied to minimize the truth table and identify combinations of conditions that are sufficient for the outcome (Ragin 1987: 937). In QCA, this is done by means of the Quine–McCluskey algorithm, also known as the truth table algorithm. It requires

Table 9.2 Example of a truth table

Conditions			Outcome	Cases
<i>X</i>	<i>Y</i>	<i>Z</i>	<i>O</i>	<i>N</i>
1	1	1	1	3
1	1	0	1	2
1	0	0	0	1
1	0	1	0	2
0	1	1	0	1
0	1	0	0	3
0	0	1	0	1
0	0	0	0	2

Source: own table.

the researcher to set a consistency threshold that determines which truth table rows will be included in the ensuing minimization procedure. By convention, this threshold should be set to a consistency of at least 0.75.

To illustrate the construction of truth tables with a simple crisp-set example, let us assume we have three conditions that are expected to cause an outcome, either in combination with other conditions or individually. Table 9.2 shows all logical combinations of conditions *X*, *Y* and *Z* and their empirical relation to the outcome *O* based on 15 cases across the eight possible configurations. The right-hand column shows us how many empirical cases fall into each configuration of conditions (i.e. truth table rows). It is apparent from the table that only the first two rows will lead towards the outcome. In Boolean algebra, this is expressed as $(X*Y*Z)+(X*Y*\sim Z)\rightarrow O$. Based on logical minimization, this complex expression can be further reduced to $(X*Y)\rightarrow O$, because the condition *Z* is irrelevant for the outcome, since *O* occurs in *Z*'s presence (see row 1) and its absence (see row 2), whereas *X* and *Y* are necessary elements of a conjunction that is sufficient for *O*. This simple example serves to illustrate the basic principle of comparison that is embedded in QCA. When fuzzy sets are used and further conditions are added, the analysis becomes increasingly complex, justifying a systematic treatment that can be conducted with the appropriate software, as will be discussed below.

Limited diversity

During the truth table analysis, most applications of QCA will encounter the phenomenon of *limited diversity*, which refers to the discrepancy between logically possible combinations of conditions (conjunctions) and the actual empirical cases in a given study. This problem increases as more conditions are included. For example, imagine a study that seeks to explain the contracting of PMSCs in 12 conflicts on the basis of five explanatory conditions. As discussed in the previous section, this would imply that there are 2^5 , or 32, distinct combinations

of conditions. Hence, because of the number of cases (conflicts) included, this design would yield *at least* 20 ‘logical remainders’ in the truth table rows not filled with empirical information; that is, if one assumes that each case shows a distinct combination of conditions (which would be a rather bold assumption). Although the phenomenon of limited diversity is ubiquitous in social research, QCA allows for analytical strategies that deal with logical remainders.³ It is therefore important for researchers to be aware of these strategies to avoid drawing unwarranted conclusions during analysis.

The QCA research process: an illustration

The remainder of this chapter will illustrate a QCA research process by drawing on my study of democratic involvement in the Iraq War (Mello 2014: Chapter 7). The study sought to explain the military participation (and non-participation) of 30 democracies in the ad hoc coalition that was established for the Iraq War in 2003. Rather than seeking a monocausal explanation for the observed phenomenon, my approach was directed towards theoretical integration, drawing on liberal, institutionalist, constructivist and neorealist explanations. The decision in favour of eclecticism was partly a response to prevalent accounts of the Iraq War, in which various factors were being highlighted as ‘causes’ but a more comprehensive explanation of military involvement was not given. Moreover, it was apparent that *interaction* between certain conditions had been overlooked, something that QCA promised to capture quite well. There are good reasons to assume that many research projects on NSAs in international security will face similar challenges and pursue comparable aims. Thus, the following reflections on the application of QCA to the study of democratic war should encourage, and is apt to guide, QCA-based research on NSAs.

Conceptualizing the outcome

For any causal research design, the first step is to specify the outcome (or dependent variable): what is to be explained? In my case, it soon became apparent that the measurement of ‘military participation’ in the Iraq War was more difficult than I had imagined when I started the project. First, there were qualitative differences in the types of military contributions. Several countries sent military engineers, others deployed ground transportation units, and some dispatched military police to Iraq. Should all these contributions be considered ‘war involvement’? Moreover, there were immense quantitative differences in the number of troops made available. For example, Estonia deployed an infantry platoon of 55 soldiers, whereas the United Kingdom sent nearly 50,000 troops to Iraq. How does one take this variation into account? Crucially, there was the dimension of time. The United States, the United Kingdom, Poland and Australia were the only countries involved in the immediate invasion in March 2003, but many other countries contributed during the following weeks and months. But where does one draw the line to distinguish military participation from other forms of involvement?

Rather than using binary coding to distinguish participation and non-participation, I decided to construct a fuzzy set, designated ‘military participation’ (MP), that allowed me to take into account the qualitative and quantitative differences in military contributions from the various countries. But I still needed a distinguishing criterion to decide whether a case was rather ‘inside’ (values larger than 0.5) or rather ‘outside’ (values lower than 0.5) the fuzzy set ‘MP’. I based this on whether or not a country’s deployment included ground troops with combat tasks – a criterion that goes to the heart of the debate about the use of military force by democracies. The number of troops and the timing of the deployments then allowed me to make further distinctions. The resulting fuzzy set MP ranged from countries such as Finland, which did not contribute (fuzzy score 0), to Belgium, which granted overflight rights (fuzzy score 0.1), to the Czech Republic, which had contributed a military field hospital (fuzzy score 0.4), to the United States, which contributed the most (fuzzy score 1.0). Table 9.3 shows the information that went into the coding of some of the selected countries (abbreviated for presentational purposes).

Selecting explanatory conditions

The selection of explanatory conditions is central to any study that uses QCA. From a theoretical point of view, the inclusion of a large number of conditions can often be desirable; however, adding further conditions will in all likelihood increase the problem of *limited diversity*, because the number of possible combinations of conditions will rise exponentially with each additional condition, resulting in fewer or no empirical cases for each row in the truth table.

In my study, like many researchers, I also faced the challenge of selecting from a potentially boundless number of factors offered by the various IR theories. I approached this problem by subsuming prevalent factors in the literature on democratic war under three broad approaches: institutional constraints, political preferences and external constraints and inducements (Mello 2014: Chapter 2). I then settled for a theoretical framework that included five explanatory conditions: (1) parliamentary veto rights, (2) constitutional restrictions, (3) executive partisanship, (4) public support and (5) military power. In line with previous work on democratic peace, which formed the backdrop to my research project, these conditions emphasized domestic factors but also included a neorealist condition based on the distribution of material capabilities among the included countries.

Of my six conditions (five explanatory conditions and the outcome), three were calibrated based on *qualitative* information by developing a coding scheme and assigning values to each case (based on country-specific research, government documents, secondary sources, news agencies, etc.). The other three conditions were coded using the ‘direct method of calibration’, which transforms raw data into calibrated fuzzy sets (on this procedure, see Ragin 2008: 85–105). For these conditions, I drew on *quantitative* data from the Comparative Manifesto Project (CMP) (Budge *et al.* 2001), various public opinion polls and military expenditure statistics.

Table 9.3 The fuzzy set ‘military participation’ in the Iraq War (outcome)

<i>Country</i>	<i>MP</i>	<i>Phase</i>	<i>Type</i>	<i>Deployment</i>	<i>Contribution</i>	<i>Troops</i>
United States	1.0	Invasion	Combat	03–2003	Army, naval, air force units	150,816
Spain	0.9	Post-invasion	Combat	04–2003	Marine infantry, support units	1300
Italy	0.8	Post-invasion	Combat	07–2003	Mechanized infantry, helicopters	2400
Czech Republic	0.4	Post-invasion	Non-combat	05–2003	Military field hospital, military police	110
Norway	0.3	Post-invasion	Non-combat	07–2003	Mine clearance	150
New Zealand	0.2	Reconstruction	Non-combat	09–2003	Engineers, reconstruction	61
Belgium	0.1	–	Logistical	–	Overflight rights	–
Finland	0.0	–	–	–	None	–

Source: own table, based on Mello (2014: 159–163).

Note

MP is the fuzzy set military participation. Table shows only selected cases for presentational purposes, for a full documentation see Mello (2014: 159–163).

Table 9.4 Fuzzy-set calibration

		<i>Austria</i>	<i>Italy</i>	<i>Japan</i>	<i>Latvia</i>	<i>Spain</i>
<i>Qualitative fuzzy sets</i>	Military participation (Outcome)	0.00	0.80	0.20	0.90	0.90
	Parliamentary veto rights	1.00	0.60	0.60	1.00	0.00
	Constitutional restrictions	1.00	0.20	1.00	0.20	0.00
<i>Quantitative fuzzy sets</i>	Public support	0.02	0.06	0.08	0.02	0.04
	Right executive	0.51	0.96	0.66	0.38	0.66
	Military power	0.15	0.63	0.87	0.13	0.25
<i>Raw data</i>	Public support (%)	8.00	18.00	20.00	7.00	12.00
	Executive L-R (CMP data)	0.80	53.83	11.12	-7.92	11.42
	Military expenditure (bn US\$)	1.80	25.60	39.50	0.15	8.70

Source: own table, based on Mello (2014, Chapter 7).

Note

Table shows only selected cases for presentational purposes, for a full documentation see Mello (2014: Chapter 7).

Table 9.4 lists the conditions of my study and shows the resulting values for five out of 30 countries. The upper three rows show fuzzy sets based on qualitative information. The lower six rows show fuzzy sets that were based on quantitative and raw data; for these data I used the direct method of calibration, after having first defined three qualitative breakpoints for each set. For the left-right (L-R) partisanship CMP data, I used breakpoints of -50 (fully out), 0 (cross-over) and 50 (fully in). Although the CMP data can take on values between -100 (all leftist statements in the party manifesto) and 100 (all rightist statements), empirical cases rarely come close to the theoretical end points of this continuum. Hence, I decided that values of 50 and -50 would be sufficient for deciding what was 'fully in' or 'fully out' of the respective sets. As a result of these coding decisions, Latvia was rather outside the set 'right executive' (0.38), whereas Italy received a value of 0.96 and could thus be considered almost 'fully in' the set of right executives.

Formulating hypotheses

QCA can be used in different ways, but most studies employ this method to test established or refined hypotheses or theories.⁴ However, because QCA investigates the specific conditions under which an outcome occurs rather than the average effect of a set of independent variables, hypotheses need to be framed in terms of necessary and sufficient conditions. This can pose a problem when the

research goal is to test established probabilistic hypotheses. Although a large body of work in the social sciences rests on a (sometimes implicit) understanding of necessary and sufficient causation (Goertz 2003b), many hypotheses in the literature continue to be framed in probabilistic language, requiring prior ‘translation’ on the part of the researcher who seeks to employ such hypotheses in an fsQCA procedure (Goertz 2003a).

In my study, I formulated two hypotheses for each condition, many of which were conceived as INUS causes (see above). Put in simple terms, this meant that the given condition was expected to cause the outcome when *combined* with other conditions. For others, the expectations derived from the literature were more concrete, so I expected a specific conjunction to be a sufficient condition for the outcome. Here are three examples (Mello 2014: 34):

- ‘Parliamentary veto rights combined with public opposition are a sufficient condition for military non-participation.’
- ‘Constitutional restrictions are a sufficient condition for military non-participation.’
- ‘Right partisanship is an INUS condition for military participation.’

Analysis of necessary conditions

The stages just described are still part of the research design; the QCA data analysis proper involves a sequence of steps, all of which can be carried out with the use of appropriate software. Until recently, the most widely used program for this purpose was fsQCA 2.5, which can perform the essential tasks but is limited in terms of reliability, advanced functions and user-friendliness. Among existing alternatives, the most promising is the QCA package for R (Thiem and Duşa, 2013), currently in version 1.1–49.⁵ In addition, a graphic user interface for R (QCAGUI) is under development and will greatly enhance the accessibility of R, especially for users who are not familiar with that environment.

The QCA procedure should always begin with a test for necessary conditions. In set-theoretic terms, a necessary condition is given when instances of the outcome are a subset of instances of a condition. As a rule of thumb, the consistency threshold for *potential* necessary conditions should be set to 0.90 (Schneider and Wagemann 2012: 143). To test for necessary conditions, I applied the formulas for consistency and coverage on each individual condition and its negation for the outcome and the non-outcome (using fsQCA for this purpose). This procedure revealed that the absence of constitutional restrictions ($\sim C$) was a necessary condition for military participation, at 0.94 consistency and 0.64 coverage. This finding supported my expectation that military participation required a lack of constitutional restrictions and, conversely, implied that constitutional restrictions amounted to a structural veto against military participation. The calculations further showed that the absence of public support was necessary for both outcomes, at 0.99 consistency for military non-participation and 0.97 consistency for military participation. Yet, given the near unanimous

public opposition to the Iraq War across the observed democracies, this finding was far from surprising. As the respective coverage scores of 0.50 and 0.55 indicated, the inferential value of this necessary condition was rather limited.

Truth table analysis

During the next step, I constructed a truth table for the outcome ‘military participation’ (Table 9.5). Because the model included five conditions, the resulting truth table comprised 2^5 (M, V, C, S, E), or 32, rows. However, because of *limited diversity*, only 12 of these rows contain empirical cases, whereas the others are logical remainders. These represent combinations of conditions that can be included in an *intermediate solution* produced by the software, if one can formulate plausible assumptions about their potential outcome even when these are not empirically observed. Table 9.5 shows the resulting truth table for military participation in the Iraq War.

Each country’s membership in the respective conjunction of conditions is given in parentheses. Italy, for instance, holds a membership of 0.60 in the conjunction given in the first row, which comprises the presence of military power, parliamentary veto rights and a right executive, combined with the absence of constitutional restrictions and public support. The consistency column indicates the extent to which the fuzzy-set values of all cases in a conjunction are sufficient for the outcome; that is, military participation. Based on the consistency scores, a cut-off point is determined for separating combinations that pass fuzzy-set sufficiency from those that do not. In my study, I elected a consistency threshold of 0.84. This meant that all configurations below Row 5 were excluded from the following minimization procedure performed by the software. I decided to exclude Row 6 because it would have lowered overall consistency but added only a single case (Norway) that held a low membership in the respective configuration (0.55). Hence, there was little inferential leverage to be gained by including this case.

The truth table is central to any QCA analysis. Even without the minimization procedure, the truth table provides comprehensive information concerning the structure of the data. As such, Table 9.5 tells us which conjunctions are filled by which empirical cases and the extent of their fuzzy-set membership. Given our theoretical expectations, we can also examine specific combinations and see whether or not these lead consistently towards the outcome.

Boolean minimization and solution terms

Based on the truth table algorithm and the consistency cut-off value specified by the researcher, the software can be used to derive three solution terms – complex, parsimonious and intermediate – that differ in the treatment of logical remainders. The *complex solution* provides a conservative estimate, because it does not make any assumptions beyond the empirical cases. As the name implies, this approach also tends to produce the lengthiest solution terms. In contrast, the *parsimonious solution* includes logical remainders but does not assess their plausibility. Although

Table 9.5 Truth table for ‘military participation’ in the Iraq War

<i>Military power</i>	<i>Parliamentary veto</i>	<i>Constitutional restrictions</i>	<i>Public support</i>	<i>Right executive</i>	<i>Military participation (outcome)</i>	<i>Consistency</i>	<i>N</i>	<i>Countries</i>
1	1	0	0	1	1	0.94	1	ITA (0.60)
1	0	0	0	1	1	0.93	2	USA (0.77), GBR (0.65)
0	0	0	0	1	1	0.87	5	AUS (0.76), ESP (0.66), NLD (0.60), PRT (0.60), POL (0.52)
0	1	0	0	1	1	0.84	4	DNK (0.84), SVK (0.60), SVN (0.58), EST (0.55)
0	1	0	0	0	1	0.84	6	LTU (0.73), BGR (0.72), ROU (0.72), HUN (0.69), CZE (0.64), LVA (0.62)
0	0	1	0	0	0	0.83	1	NOR (0.55)
1	0	0	0	0	0	0.77	1	FRA (0.68)
1	1	1	0	1	0	0.65	1	JPN (0.60)
0	1	1	0	1	0	0.58	1	FIN (0.60), AUT (0.51)
1	1	1	0	0	0	0.58	1	DEU (0.73)
0	0	0	0	0	0	0.54	4	NZL (0.80), GRC (0.78), CAN (0.76), BEL (0.71)
0	1	1	0	0	0	0.45	2	SWE (0.82), IRL (0.75)

Source: own table, based on Mello (2014: 172).

this leads to solution terms that are easier to interpret, the parsimonious solution should always be treated with care and contrasted with other solutions, because it could comprise implausible assumptions. Finally, the *intermediate solution* allows the researcher to designate how logical remainders are to be treated based on explicit expectations about the causal relationship; hence, it is positioned between the complex and the parsimonious solutions.

Table 9.6 shows the analytical results for war involvement in Iraq. The top rows display the previously identified necessary conditions. As expected from the analysis of necessity, all paths towards military participation contained the absence of constitutional restrictions ($\sim C$). The complex solution further entailed public opposition ($\sim S$), which is also implicated in the other solution terms but is not part of the minimized formulas. Measures of consistency and raw coverage are given for each solution term, whereas solution paths further specify raw coverage and unique coverage. The latter indicates 'how much' is explained *exclusively* by a respective path, whereas raw coverage also includes empirical overlap. For example, the parsimonious solution term includes two paths and has an overall consistency of 0.77 and coverage of 0.85; however, we can see that Path 1 has a consistency of 0.82 and a *unique coverage* of 0.28 (indicating how many cases this path can account for).

The solution terms show that two consistent pathways towards military participation exist. The first entails a right executive (E) and the absence of constitutional restrictions ($\sim C$), as indicated in Path 3. The second comprises the absence of

Table 9.6 Solution terms for 'military participation' in the Iraq War

	<i>Conjunction</i>	<i>Relation</i>	<i>Consistency</i>	<i>Raw coverage</i>	<i>Unique coverage</i>
Necessary condition					
	$\sim C$	$\leftarrow MP$	0.94	0.64	–
	$\sim S$		0.97	0.50	–
Parsimonious solution term			0.77	0.85	–
Path 1	$\sim C * E$	+	0.82	0.69	0.28
Path 2	$V * \sim C$	$\rightarrow MP$	0.75	0.57	0.17
Intermediate solution term			0.79	0.85	–
Path 3	$\sim C * E$	+	0.82	0.69	0.31
Path 4	$\sim M * V * \sim C$	$\rightarrow MP$	0.77	0.54	0.17
Complex solution term			0.80	0.84	–
Path 5	$\sim C * \sim S * E$	+	0.84	0.68	0.30
Path 6	$\sim M * V * \sim C * \sim S$	$\rightarrow MP$	0.78	0.54	0.17

Source: own table, based on Mello (2014: 173)

Note

[M] Military Power, [V] Parliamentary Veto, [C] Constitutional Restrictions, [S] Public Support, [E] Right Executive, [MP] Military Participation, [\sim] absence of a condition, [*] logical 'and', [+] logical 'or', [\leftarrow] necessity, [\rightarrow] sufficiency.

military power ($\sim M$) and constitutional restrictions ($\sim C$) with parliamentary veto rights (V), as shown in Path 4. This finding lent some support to the partisan argument and the expectation that small powers without constitutional restrictions had had incentives to become involved militarily in the Iraq War.

Visualization of results

To assess the validity of the analytical results and to relate them to individual cases, graphic means of representation can be helpful. For this purpose, I constructed an x - y plot that displayed each country's membership in the complex solution term in relation to its membership in the outcome. The plot showed that the complex solution was (almost) sufficient for military participation because a large majority of countries were placed above the main diagonal, which reflects the subset relationship. As defined above, a sufficient condition X is given whenever X resembles a *subset* of Y , meaning that all values for X are equal to or less than the values for Y . Hence, cases on or above the diagonal fulfil the subset criterion for a sufficient condition.

Figure 9.1 shows four groups of cases. Countries in the lower-left corner hold low membership values in outcome and solution and can thus be considered

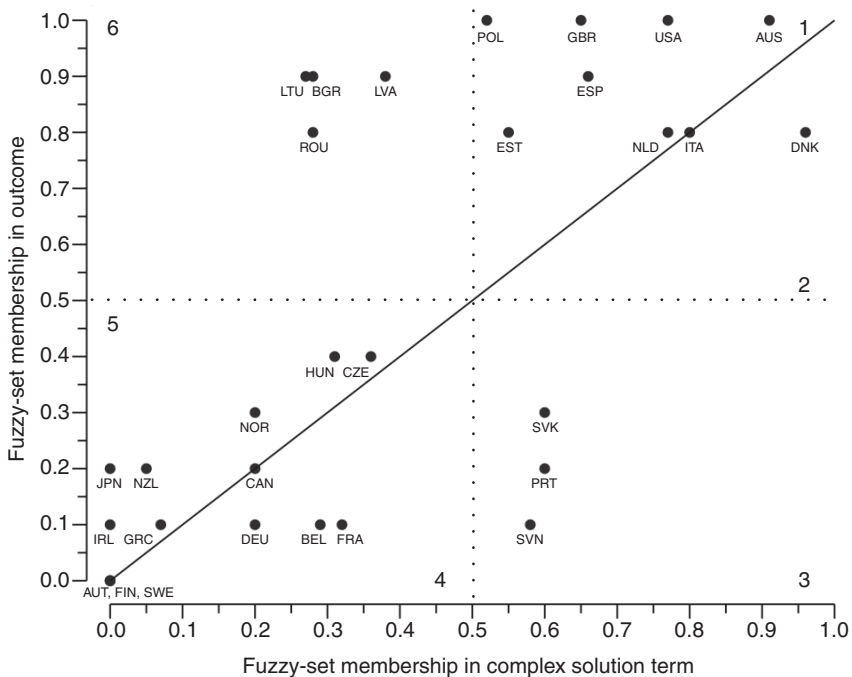


Figure 9.1 Military participation and solution term.

Source: own figure, based on Mello 2014: 174.

largely irrelevant for the analysis. In contrast, 12 out of 30 democracies hold membership in the solution term (Zones 1 to 3), eight of which can be considered typical cases (Zone 1). In turn, Zone 3 holds three deviant cases, as in countries with membership in the solution that do not show the expected outcome. In other words, based on their characteristics, one would have expected Portugal, Slovakia and Slovenia to contribute to the Iraq War beyond their actual involvement. Finally, Zone 4 includes countries that show the outcome but do not hold membership in the solution. This indicates that alternative explanations may better account for the pattern observed in these particular cases.

Depending on the specific aim of a research project, the next step could be the selection of cases for an in-depth study by means of process-tracing (see Chapter 8 by Andreas Kruck). Here, the x - y plot can be useful in identifying appropriate cases for further study (Schneider and Rohlfing 2013). For example, in Figure 9.1, we could select a case either from Zone 3 to explain why the country did not participate as expected or from Zone 6 to discuss alternative explanations for the observed participation that were not covered by our theoretical framework.

Interpretation of results

QCA solution terms tend to be complex and difficult to interpret. Hence, it is important to formulate specific theoretical expectations *before* the analysis and relate these to the results provided by the software. Rather than focusing solely on the analysis of the outcome, it is further recommended to also examine the *non-outcome*. That being said, for reasons of space I will solely highlight some selected findings for the analysis of the outcome of military participation (for details, see Mello 2014: 176–181).

First, the absence of constitutional restrictions was found to be a necessary condition for military participation and was also part of all sufficient conjunctions. This finding contradicted the argument presented in some previous studies that constitutional settings would not constrain decision-making regarding the use of force. Second, the analysis shed light on the relationship between partisanship and war involvement. As expected, left or right partisanship were *individually* neither necessary nor sufficient. Instead, it was confirmed that these constitute INUS conditions. A rightist government combined with the absence of constitutional restrictions was sufficient for involvement in the Iraq War. Countries with a high level of membership in this path included, among others, Australia, Spain and the United States. Finally, the ‘parliamentary peace’ hypothesis could not be confirmed. Earlier studies suggested that parliamentary veto rights should serve as an effective constraint against war involvement (Dieterich *et al.* 2015), especially when combined with widespread public opposition to the use of force ($V^* \sim S$). Although this pattern was found empirically, it crossed with constitutional restrictions and was thus overdetermined with regard to explaining military non-participation. Moreover, countries such as Italy, Denmark, Bulgaria, and several others also held membership in ($V^* \sim S$), but for them the expected mechanism failed to prevent military involvement.

Conclusion: strengths and limitations of QCA

QCA has gained recognition as a research approach that offers distinct advantages for comparative studies. Although many existing QCA applications take a *macro-comparative* perspective, there is nothing inherent in the set-theoretic approach that would prevent it from being used for the study of NSAs in the field of international security. Hence, this chapter should also be understood as a plea for the (wider) application of QCA in this specific area of research. To this effect, it has provided an outline of the core principles and terminology of QCA and complemented this methodological introduction with a step-by-step illustration of an empirical study.

As discussed in the theoretical section and shown in the empirical part of this chapter, QCA allows the researcher to take into account *causal complexity* – most importantly, the fact that different pathways can lead towards the same outcome and that a combination of conditions can be jointly necessary and/or sufficient for an outcome. Moreover, QCA offers a *systematic* and *rigorous comparative approach*. Yet it gives researchers *flexibility*, because crisp and fuzzy sets can be calibrated on the basis of qualitative and quantitative sources of data, depending on the particular research aim and the type of information required to operationalize one's theoretical framework.

Despite these strengths, it should be noted that QCA also has several limitations.⁶ First, a researcher interested in using QCA would need to 'buy into' the set-theoretic logic, because it is the central methodological assumption of the approach. Certainly, it would make no sense to test probabilistic hypotheses within a set-theoretic framework. Likewise, if the language of necessary and sufficient causation does not fit with the research aim of a given study, then QCA should simply not be used. Second, the truth table analysis itself is a static approach. Hence, it is difficult to incorporate assumptions about sequence and timing in a QCA framework because the analysis treats all conditions the same way. There are variants of QCA that try to circumvent this limitation and include 'time' as a factor; but as of now, these still come with substantial drawbacks (see Schneider and Wagemann, 2012: 263–274). Finally, in order for QCA to 'get off the ground' and work properly, a certain number of cases are required, because it gets increasingly difficult to conduct the truth table analysis if there are not enough observations. Although some studies have involved as few as nine cases, this numerical restriction might be the biggest obstacle to using QCA for some research projects on NSAs. Nonetheless, when the requirements of a set-theoretic approach are met, the application of QCA can certainly be a rewarding strategy for research projects on NSAs in international security.

Notes

- 1 For a demonstration of the effects of different calibration techniques and their application in QCA, see Ragin (2008: 85–105) and Schneider and Wagemann (2012: 32–41).
- 2 The mathematical formulas behind these measures are introduced and discussed in Ragin (2006).

- 3 For an elaborate discussion of these strategies, see Schneider and Wagemann (2012: Chapter 6).
- 4 On various ways to use QCA, see Berg-Schlosser *et al.* (2009: 15–16).
- 5 For a comparison of the various software packages, see Schneider and Wagemann (2012: 283).
- 6 For a comprehensive discussion of the strengths and limitations of QCA, see the symposia published in the APSA Newsletter *Qualitative Methods*, 2004 (2)2, in *Studies in Comparative International Development*, Spring 2005 (40)1, *Political Research Quarterly* 2013 (66)1 and *Sociological Methodology* 2014 (44)1. A detailed reply to some of the critiques can be found in Meur *et al.* (2009).

References

- Berg-Schlosser, Dirk, Gisèle De Meur, Benoît Rihoux, and Charles C. Ragin. 2009. 'Qualitative Comparative Analysis (QCA) As an Approach', pp. 1–18 in *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, B. Rihoux and C. C. Ragin, eds. Thousand Oaks, CA: Sage.
- Blatter, Joachim, and Markus Haverland. 2012. *Designing Case Studies: Explanatory Approaches in Small-N Research*. Houndmills: Palgrave Macmillan.
- Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, and Eric Tanenbaum. 2001. *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments, 1945–1998*. Oxford, UK: Oxford University Press.
- Collier, David, Henry E. Brady, and Jason Seawright. 2010. 'Critiques, Responses, and Trade-Offs: Drawing Together the Debate', pp. 135–159 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edn. H.E. Brady and D. Collier, eds. Lanham: Rowman & Littlefield.
- Dieterich, Sandra, Hartwig Hummel, and Stefan Marschall. 2015. 'Bringing Democracy Back In: The Democratic Peace, Parliamentary War Powers and European Participation in the 2003 Iraq War', *Cooperation and Conflict* 50(1): 87–106.
- Gerring, John. 2012. *Social Science Methodology: A Unified Framework*. 2nd edn. Cambridge: Cambridge University Press.
- Goertz, Gary. 2003a. 'Cause, Correlation, and Necessary Conditions', pp. 47–64 in *Necessary Conditions: Theory, Methodology, and Applications*, edited by G. Goertz and H. Starr, eds. Lanham: Rowman & Littlefield.
- Goertz, Gary. 2003b. 'The Substantive Importance of Necessary Condition Hypotheses', pp. 65–94 in *Necessary Conditions: Theory, Methodology, and Applications*, G. Goertz and H. Starr, eds. Lanham, MD: Rowman & Littlefield.
- Mackie, John L. 1965. 'Causes and Conditions'. *American Philosophical Quarterly* 2(4): 245–264.
- Mahoney, James. 2010. 'After KKV: The New Methodology of Qualitative Research'. *World Politics* 62(1): 120–147.
- Mahoney, James, Erin Kimball, and Kendra L. Koivu. 2009. 'The Logic of Historical Explanation in the Social Sciences'. *Comparative Political Studies* 42(1): 114–146.
- Mello, Patrick A. 2012. 'Parliamentary Peace or Partisan Politics? Democracies' Participation in the Iraq War'. *Journal of International Relations and Development* 15(3): 420–453.
- Mello, Patrick A. 2013. *From Prospect to Practice: A Critical Review of Applications in Fuzzy-Set Qualitative Comparative Analysis*. Paper prepared for presentation at the Eighth Pan-European Conference on International Relations, Warsaw, 18–21 September 2013.

- Mello, Patrick A. 2014. *Democratic Participation in Armed Conflict: Military Involvement in Kosovo, Afghanistan, and Iraq*. Houndmills: Palgrave Macmillan.
- Meur, Gisèle De, Benoît Rihoux, and Sakura Yamasaki. 2009. 'Addressing the Critiques of QCA', pp. 147–165 in *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, B. Rihoux and C.C. Ragin, eds. Thousand Oaks: Sage.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press.
- Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago, IL: University of Chicago Press.
- Ragin, Charles C. 2006. 'Set Relations in Social Research: Evaluating Their Consistency and Coverage'. *Political Analysis* 14(3): 291–310.
- Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.
- Rohlfing, Ingo. 2012. *Case Studies and Causal Inference: An Integrative Framework*. Houndmills: Palgrave Macmillan.
- Schneider, Carsten Q., and Ingo Rohlfing. 2013. 'Combining QCA and Process Tracing in Set-Theoretic Multi-Method Research'. *Sociological Methods & Research* 42(4): 559–597.
- Schneider, Carsten Q., and Claudius Wagemann. 2010. 'Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets'. *Comparative Sociology* 9(3): 397–418.
- Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge, UK: Cambridge University Press.
- Thiem, Alrik, and Adrian Duşa. 2013. *Qualitative Comparative Analysis with R: A User's Guide*. New York: Springer.
- Zadeh, Lotfi A. 1965. 'Fuzzy Sets.' *Information and Control* 8(3): 338–353.

10 Geospatial analyses of non-state actors in violent conflicts

Alexander De Juan

Introduction

This chapter focuses on the use of quantitative geospatial analysis for investigating non-state actors (NSA) in violent conflicts. How can information on the geographical associations between socio-economic factors and violence be used to answer questions related to the motives, opportunities and strategies of militias and rebel groups? Drawing on my own research concerning communal violence in the Darfur region of Sudan, I illustrate how this specific methodological approach can be translated into a concrete research design, and how problems that arise during the research process can be dealt with. In particular, I will focus on conceptual issues related to the generation of plausible spatial hypotheses, the selection of units of analysis, the operationalization of core variables and the interpretation of results.

When anti-state opposition escalated in Darfur in 2003, the government enlisted the help of militias. These mostly Arab proxy armies sided with the government against mainly non-Arab groups (Salmon 2007). Media and policy-related accounts of the conflict have argued that climate change and resource scarcity were essential in motivating militia attacks and the ensuing inter-ethnic violence. I have analysed whether longer-term changes in the subnational distribution of water and vegetation resources really correlated with the spatial distribution of militia violence in the early phase of the Sudanese civil war. Investigating highly disaggregated geographical patterns of violence and decadal environmental change in an information-scarce region such as Darfur requires a specific research approach. The core empirical analysis of the study consisted of a geospatial quantitative analysis based on remote sensing information about attacks on villages from 2003 to 2005 and on changes in vegetation in the 20 years prior to the beginning of the civil war. By complementing these data with additional socio-economic information, such as the spatial distribution of the population and access to infrastructure, I was able to construct a data set of artificial grid cells covering all populated areas of northern and southern Darfur. More specifically, using the open-source Geographic Information System (GIS) software Quantum GIS (QGIS), I projected a grid net over the region of interest and used individual grid net cells as units of analysis, measuring all variables of

interest per individual cell. Statistical analyses using Stata 12 have yielded robust correlations between the probability and intensity of violence and long-term changes in natural resource availability (see De Juan 2015).

In the following sections, I describe the core methodological and conceptual decisions that led me to this specific research design. First, I outline some major arguments in favour of subnational geospatial analyses. I then address individual elements of the research process. Throughout the subsections, I present more general considerations, as well as practical descriptions of how I approached the respective issues in my own research. The final section summarizes key methodological issues and recommendations.

Subnational geospatial analyses of NSA

In peace and conflict studies, countries constitute the prime units of analysis. Indeed, various factors relevant to understanding the motives, strategies and activities of rebel groups are best analysed at the country level, either because these factors do not display substantial within-country variations (e.g. regime type) or, as one might argue, because aggregate country-level features are more important than variations on more disaggregated levels (e.g. ethnic or religious demography). However, many other aspects of NSA may actually be better analysed with alternative units of observation. A micro-level turn has reoriented research on political violence from structural factors at the country level to disaggregated analyses of dynamics and actors at a subnational level (Kalyvas 2006).

As part of this more general trend, quantitative geospatial analyses have been gaining in importance (e.g. Buhaug and Gates 2002; Kalyvas 2008; Raleigh *et al.* 2010a, 2010b). These studies use quantitative techniques to investigate correlations between outcomes and explanatory factors of interest across geographically disaggregated units of analysis, such as regions, districts or artificial grid cells. As such, they use information on the ‘Where’ to answer questions related to the ‘Why’ of armed conflicts: Why do rebel groups emerge? Why do people support violent action? Why do rebel groups act in a specific way? Thus, geographical variations are not the actual research interest, but are rather an instrument for inference. For instance, based on the fact that violence is geographically concentrated in a particularly poor region of a country, we might infer that economic deprivation plays a key role in motivating rebellion against the respective state. Two key advantages of this methodological approach stand out: it avoids the ecological inference fallacy, and it considers the spatial characteristics and determinants of NSA activities.

The ecological inference fallacy can occur when aggregated data are used to draw conclusions about individual characteristics or relationships (Robinson 1950; King 1997). Most NSA are not active across entire state territories; their presence and activity are geographically limited to specific substate regions. The same holds true for factors that are said to influence rebel group emergence and activity, such as ethnic identity, terrain or economic well-being. For example, a comparatively high national gross domestic product (GDP) may indicate a high

degree of welfare at the national level. However, national-level numbers may obscure the strong concentration of economic power in specific regions, which in turn may exacerbate regional, ethnic or religious cleavages and increase rather than reduce rebel groups' potential for mobilization. Without analysing whether subnational patterns of violence match the subnational patterns of the explanatory variables, it is difficult to draw reliable conclusions concerning the actual effects of the factors under investigation (see for example Raleigh *et al.* 2010b; Raleigh 2011).

Subnational geospatial analyses allow not only new ways of analysing conventional questions related to rebel groups and other NSA but also the analysis of new questions. Spatial disaggregation permits the consideration of characteristics of NSA that are very difficult or impossible to investigate on the national level. For example, studies may analyse the spatial diffusion of NSA activities, specific features of rebel governance or how the same NSA behave differently across different subnational regions. The subnational approach also allows the consideration of specific explanatory variables that are inherently geographical, including economic and political disparities across regions (Fjelde and Østby 2012), peripheral locations (see for example Rustad *et al.* 2011) and the accessibility of regions (Raleigh and Hegre 2009). This diversification of explanatory and outcome variables may help us understand rebel groups and other NSA more comprehensively and to scrutinize aspects that have previously been under-investigated.

Designing geospatial analyses

Geospatial studies not only have specific potentials but also face specific methodological challenges. A number of studies have dealt in detail with more technical issues related to the statistical analysis of spatial data (see for example Gleditsch and Weidmann 2012; Raleigh *et al.* 2010b; Stephenne *et al.* 2009).¹ The following subsections focus on conceptual specificities of geospatial analysis, the kinds of problems they may create and how they may be dealt with. This section is organized according to essential elements of the research process in which these challenges may occur: (1) the definition of the theoretical argument; (2) the selection of the unit of analysis; (3/4) the conceptualization and measurement of the outcome and explanatory variables; and (5) the meaningful interpretation of quantitative results.

Suggesting plausible hypotheses

In geospatial analysis, defining a plausible causal hypothesis is more challenging than it may seem at first glance. Geospatial analyses differ from country-level studies in at least one essential respect: whereas the presence of rebel groups is presumably associated with structures, actors and processes observable within the same country, this is not necessarily the case at the subnational level (Kalyvas 2008). People do not necessarily fight in the subnational region where

their grievances or personal political or economic aspirations have impelled them to take up arms. Thus, in contrast to the country-level approach, geospatial analysis does not allow us to establish a direct link between the location of NSA activity and the motives of the NSA. The following blunt example may illustrate this argument.

Subnational analyses may reveal that rebel groups are particularly active near military bases or government facilities. The explanation is straightforward: challengers of state authority act strategically and target the resources of their enemies. Thus, the location of NSA activity informs us about the strategic aims of the rebels but not – or not necessarily – about their actual motives. Similarly, in many cases, NSA activities within a subnational region may be a consequence of diffusion from neighbouring regions rather than being associated with specific characteristics of the region itself (Schutte and Weidmann 2011). Rebellions erupt in one region and may in time engulf other parts of a country. If we observe violent events within a given subnational unit of analysis, we therefore cannot know for sure whether violence has been caused by factors found in that unit, whether violence has been caused in other areas but has been ‘pulled’ to that unit as a result of specific characteristics of that unit (e.g. features of terrain) or whether violence has been ‘pushed’ to that unit from surrounding areas (e.g. as a consequence of a military retreat).

It is for this reason that one must carefully translate basic assumptions about the NSA under investigation into concrete hypotheses based on observable spatial associations. Specifically, one needs to develop explicit auxiliary theories that link the location of explanatory factors to the location of observable NSA activities. In the case of my study of Darfur, careful reflection about the causal steps that linked environmental change to NSA activity led to seemingly counter-intuitive hypotheses. Although one might at first assume that violent conflict would grow more intense where natural resources were becoming scarcer, a closer look at the context made another hypothesis more plausible. Previous arguments had highlighted environmentally induced migration to relatively resource-rich areas as an essential mechanism that links political violence to the more long-standing processes of environmental change (Reuveny 2007). Such change was uneven in Darfur: in the northern and eastern parts of the region, natural resources became depleted, whereas in the south and west, resources became more available. These protracted processes eventually resulted in migration to the more affluent south and west. This led me to the hypothesis that violent acts against civilians would become more likely and more intense in those areas where the influx of migrants would increase competition for resources among identity groups that lacked common traditions and institutions of peaceful conflict resolution.

I then developed an explicit auxiliary hypothesis that translated my general (aggregated) hypothesis – increasing resource scarcity and competition motivated inter-ethnic militia attacks – into a concrete spatial argument: resource scarcity in resource-poor areas led to high levels of in-migration and resource competition in relatively resource-rich areas and increased the risk and intensity of militia attacks in these regions.

Selecting the 'right' geographical unit of analysis

At what geographical scale do I expect these hypothesized causal associations to materialize? Switching from the country-level to the subnational level of analysis opened up choices as to what geographical unit of analysis should be selected to investigate a specific research question: should the explanatory and outcome variables be measured at the level of villages, districts or provinces, or should one create a small or large artificial grid net that covers the geographical area under investigation and then measure the variables per grid cell? For example, if I wanted to investigate political violence in Pakistan, I could use one of the following four units of observation (see Figure 10.1): larger or smaller administrative units (e.g. districts or subdistricts) or two different-sized grid cells (e.g. individual square units of, say, $100\text{ km} \times 100\text{ km}$ or $50\text{ km} \times 50\text{ km}$, which together make up grid nets).

Although these options increase the freedom of the researcher, they present the challenge of having to make and justify an actual decision. Here the issue is what units of analysis are appropriate for a specific research project (Buhaug and Rød 2006; Rød and Buhaug 2007; De Luca *et al.* 2012). In political geography, the widely debated modifiable areal unit problem (MAUP) refers to potential variation of results across different units of analysis (Openshaw 1996; O'Loughlin 2007; Lee *et al.* 2008). The MAUP basically consists of two different aspects. One of these aspects, zoning, concerns the question of how geographical boundaries of units of observation should be defined. In most cases, researchers do not know what determined the shapes of provinces or districts, such as geographical characteristics or social criteria (e.g. ethnic or religious factors). Thus, it is not clear what these boundaries really signify and how they relate to the social phenomena under investigation (Openshaw 1996). The other aspect of the MAUP is scaling, which concerns questions such as: in geographical terms, how small or large should the units of observation be; should specific phenomena be analysed at the village level or at the district level; and what is the optimal resolution of the grid cells to be used?

Conceptual challenges may arise when one is selecting the unit of observation. For example, with respect to the relational variables often used in analyses



Figure 10.1 Shape and distribution of four different units of analysis used to analyse violence in Pakistan.

Source: own figure.

Note

From left to right: Districts, subdistricts, $100\text{ km} \times 100\text{ km}$ grid cells and $50\text{ km} \times 50\text{ km}$ grid cells.

of political violence, such as vertical or horizontal inequality and ethnic and religious polarization or fractionalization, one could argue that at the national level people will tend to compare their own situation with that of fellow citizens of their country. However, this argument cannot easily be made at the subnational level. Does inequality matter across or within districts, and across or within villages? How should grid cells be positioned and sized in order to capture relevant demographic (religious or ethnic) constellations?

There is no such thing as an optimal size and shape of units of analysis in investigations of specific NSA features and activities. Smaller units increase the chance of maximizing the homogeneity within and the variations across the units of observation (Oberwittler and Wikström 2009). Also, if one assumes that people's perceptions and actions are shaped by their immediate local environment rather than by wider administrative units or grid cells (Lee *et al.* 2008), smaller units might be more suitable. Some theoretical arguments to be found in studies of peace and conflict, however, emphasize the role of factors at the level of larger units of analysis. For example, if one assumes that the ethnic composition of decentralized regional governments plays a role in political violence, it would not be reasonable to investigate this argument using more disaggregated units of analysis, such as districts or villages. Moreover, it may make sense to use larger units of observation as a way to account for low-level geographical gaps between the location of the cause and the location of the observable effect (Raleigh and Kniveton 2012). For instance, if we expect greater violence near military bases because rebels strategically target critical infrastructures, we may expect violent events to be concentrated near these bases but not necessarily at their exact geographical location.

Similarly, certain arguments favour the use of grid cells rather than the use of administrative units, or vice versa. For example, grid cells may be more favourable because, unlike administrative units such as districts, they are not intentionally shaped based on economic, political or social criteria unknown to the researcher (O'Loughlin 2007). Moreover, grid nets can more easily be generated in different sizes and locations. From a practical perspective, grid cells may be easier to handle in cross-country and panel analyses, because administrative areas often vary greatly in size across countries, and their boundaries are often subject to amendments. On the other hand, administrative units may be more suitable when reliable socio-economic data are obtained at the administrative level (e.g. from household surveys) rather than at less aggregated levels (e.g. per household or per individual), which would be required for aggregation into artificially created grid cells.

Decisions concerning units of analysis should be made on a case-by-case basis in line with the specific research questions and factors under investigation (Openshaw 1996; Lee *et al.* 2008). These decisions must be explicit and must be validated by conceptual arguments in favour of their selection. If the units of observation are selected purely on the basis of data availability, researchers should explain why this specific unit of analysis is suitable for the respective analyses and how zoning or scaling effects may affect the results. Whenever

possible, robustness checks using a variety of different units of analysis should be conducted to show that the results are not driven by a more or less arbitrary decision to use a specific unit of analysis.

In my study of Darfur, I relied on artificial grid cells as the units of observation. Administrative units were just not a practicable choice, because the principal administrative units – states and counties – are fairly large in Darfur, so that a lot of geographical variation of interest would have been concealed by aggregation into these units. Moreover, they promised few advantages in terms of data availability, because administrative data for Sudan are scarce. I therefore relied on a grid net that I modelled according to previous extensive environmental and livelihoods vulnerability mapping (ELVM) that was undertaken in Darfur by the International Organization of Migration (IOM) and the ProAct Network (2010). Modelling the units of analysis according to an existing study had certain disadvantages: it limited the possibility of performing robustness checks with alternative grid specifications, most notably in terms of size and location, and I had to focus on the specific study areas of the assessment, which excluded western Darfur.

The essential advantage of this approach was that it allowed me to use the data provided by the study on the respective grid cell level. In a context of scarce socio-economic data, this advantage overruled other considerations and led to my decision to use a total of 1087 grid cells, 16 km × 16 km each, across northern and southern Darfur as my primary units of analysis. To control for the MAUP, I also created a second data set on the village level that was used for additional robustness checks. Instead of aggregating data per 16 km × 16 km grid cell, I measured them in circular village buffers with a radius of either 2 km or 5 km. Statistical analyses with these alternative units of analysis confirmed my initial findings and made me confident that they were not driven by the selection of my primary unit of analysis.

Measuring the outcome variables

The next step in the research process consisted of measuring the main variables of interest. For my study of Darfur, I needed geolocated information on militia activities – but where could such data be obtained?

Most subnational analyses of various facets of violence rely on newspaper-based event data sets such as the Armed Conflict Location & Event Data Project (ACLED) or the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) (see Raleigh *et al.* 2010a; Sundberg and Melander 2013). These data sets, which rely on international press coverage to identify violent events that are in turn attributed to geographical locations described in the respective articles, have provided extremely valuable avenues for research on geographical patterns of violence. Their specific strength is that they provide temporally and geographically disaggregated data that have been collated based on the sources and coding rules used by a number of countries, thus allowing analyses of violence on the subnational level across several countries.

One weakness of these data sets, however, is that the geographical information is not always precise or complete, and because they depend on international press coverage, violence in remote and rural areas is probably under-reported. As has been emphasized elsewhere, these data sets also do not always correctly attribute the place names cited in press reports to specific locations (Eck 2012). These issues raise problems for analyses on geographically highly disaggregated levels, as in the case of geospatial analyses of single countries or regions. These data-collection challenges are likely to be more pronounced in regions where violence is rampant in inaccessible places and place names are poorly documented.

In many countries, more specific data sets that rely on national newspaper sources are available. In Nigeria, for example, a data set of violent events was compiled by the non-profit organization Nigeria Watch, which screened 15 local newspapers and human rights reports from 2006 to 2012 to extract information on 9255 violent events. The BFRS Political Violence in Pakistan Dataset (Bueno de Mesquita *et al.* 2013) contains incident-level data on political violence that took place in Pakistan between 1988 and 2011, with a total of 17,679 events and rich information concerning the state and non-state actors involved. The Sistem Nasional Pemantauan Kekerasan Indonesia (SNPKI) is a government- and World Bank-funded violent event database that covers several of Indonesia's provinces; begun in 1998, this project provides a monthly updated list of violent events, which now exceeds 119,107 incidents.

More recently, a number of crowdsourcing platforms that monitor violent events have been established. Such initiatives can provide abundant information even where international and national press coverage is minimal (see for example De Juan and Bank 2015). Individuals use online resources to file reports that either describe specific locations or position them on digital maps. Finding such country-specific data sets is often difficult and requires rigorous research concerning the countries of interest. Because most of these data sets are less prone to suffering from the problems associated with the cross-country data sets noted above, this additional investment will often pay off by providing greater data reliability. This is not to say that country-specific data sets are free from biases. Crowdsourcing data in Syria, for example, will likely be biased towards the opposition. Regions that are aligned with the ruling regime may tend to under-report violence, whereas rebel strongholds will most likely over-report numbers of fatalities. In addition, certain types of incidents may not be accurately reflected in the data; for example, particularly intense fighting may be under-reported if fewer civilian observers are able to file reports on such incidents.

In any case, it is essential to make an informed choice about what data should be used while acknowledging data set-specific drawbacks and challenges. As for my study on Darfur, I doubted the reliability of ACLED and GED data, because these data sets were probably beset by such challenges in the sparsely populated Darfur region, which had little international press presence and coverage. Consequently, I tried to find alternative country-specific data. In order to document

violence in Darfur, the Humanitarian Information Unit (HIU) at the US State Department created a database of information about villages that were damaged (evidence of partial destruction) or destroyed (evidence of complete destruction) between 2003 and 2010, as well as about those confirmed to be neither damaged nor destroyed during this period. Derived from satellite images that revealed changes in village structures over time, the data set contains information on 9183 villages across Darfur.

Once a data set has been selected, the next step involves operationalizing the outcome variable according to the main theoretical arguments. Event data may be used to measure various facets of rebel group violence; that is, its onset, incidence or intensity. Although the determinants of all three of these elements may overlap, they are not necessarily the same (Montalvo and Reynal-Querol 2007). The factors that motivate people to take up arms or that influence their ability to stage a rebellion may differ from the factors that influence the intensity of fighting. A large number of events in a specific area may be traced back to a spiral of attacks and reprisals within a given territory (Raleigh 2012), or it may be influenced by a military scorched-earth strategy aimed at weakening the adversary. In neither case are the factors that increase the intensity of the conflict within a specific area necessarily related to the causes of the conflict. The same argument can be made the other way around: a high degree of poverty in a subnational region may be associated with a greater risk or incidence of conflict within that region (Buhaug *et al.* 2011); however, why high levels of poverty should actually increase the intensity of fighting within a region is not theoretically obvious, because having compelling reasons to fight does not necessarily lead to more zealous, aggressive fighting (see for example Costalli and Moro 2011).

My principal theoretical argument concerning the role of environmental change in non-state violence in Darfur explicitly applied to the occurrence and intensity of militia violence; that is, inter-ethnic competition in resource-rich areas increased the likelihood of attacks on villages. Resource wealth also creates incentives for high-intensity violence aimed at displacing competing ethnic groups. Consequently, I decided to consider both variables in the main models of my analysis. The first was a dichotomous ‘occurrence variable’, which was assigned a value of 1 for each grid cell that contained at least one village found to be damaged or destroyed during the period under investigation, and a value of 0 if no villages in the grid cell were damaged or destroyed. The second was an ‘intensity variable’, which measured the actual counts of villages damaged and/or destroyed per grid cell. The results were comparable for both specifications, indicating that environmental factors played an important role in the occurrence and intensity of militia violence.

Measuring the explanatory variables

Regardless of the level of analysis and the specific methodological approach, it is often difficult to identify indicators of the explanatory variables of interest that are both valid and reliable. Most classical theoretical arguments concerning the

causes of political violence focus on individual needs, strategies and actions; for example, the role of deprivation in generating frustrations that may motivate individuals to take up arms (Gurr 1970). Thus, all causal theories about conflict onset, intensity or duration entail explicit or implicit assumptions regarding the effects of the respective explanatory factors on individual-level perceptions, decisions and actions. The use of aggregated indicators (e.g. country-level or district-level indicators) to explain political violence requires auxiliary assumptions that link the indicator to local, micro-level dynamics; however, such assumptions introduce additional uncertainties into empirical analyses. We might assume, for example, that a state's security budget reflects its actual coercive capacity on the ground, or that decentralization laws mirror individuals' actual access to decentralized decision-making.

A major challenge confronting every quantitative study is the need to translate associated abstract concepts and variables into quantifiable and measurable indicators suitable for more aggregated statistical analysis (Kalyvas 2003, 2008). Lack of data can make persuasive operationalization in subnational analyses even more difficult, especially in conflict-ridden countries such as Sudan. Whereas information about the national-level GDP may be available, information about income levels at the subnational level is often unattainable. The same holds true for data about ethnic identities or various development indicators. Paucity of data increases the need to rely on proxy indicators that may or may not appropriately reflect the concepts at hand. Thus, geospatial analyses have, for example, attempted to capture state capacity using such indicators as distance from the capital, type of terrain (e.g. mountainous) and density of road networks (Cederman *et al.* 2009; Buhaug 2010).

Another problem is that individual indicators may in principle be used as proxies for various variables and concepts, thus making the interpretation of empirical results somewhat discretionary and ambiguous. For example, if we consider road networks, in addition to reflecting state capacity (Cederman *et al.* 2009), they might be used to indicate overall levels of development (Buhaug and Rød 2006), or they might be seen as strategic targets of rebel groups (Raleigh and Hegre 2009) or as potentially enhancing rebels' logistical capacities (Zhukov 2012). What conclusions can be drawn if dense road networks are found to be negatively or positively associated with political violence?

Challenges of measurement related to a lack of available data are not easily resolved. One way of dealing with such challenges is to focus on countries with a relative abundance of relevant data, although this strategy is not very satisfying. It excludes from the analyses many countries in which violent NSA play a particularly pronounced role, most notably fragile states with high levels of violence and low statistical capacity and/or a high level of political interference with the generation of statistical data. Thus, a better strategy for dealing with a lack of reliable data consists of more rigorously searching for less obvious data sources. Various data sets contain geographically disaggregated information, such as the US Agency for International Development's Demographic and Health Surveys (DHS) and EdData (Education Data); the World Health Organization's World

Health Survey (WHS); and the World Bank's Living Standards Measurement Study (LSMS). In many cases, such data feature information (on identities and attitudes, as well as on the social and economic conditions of households and individuals) that is relevant to the analysis of NSA.

Remote sensing data can provide relevant geographical information on a highly disaggregated level. Satellite images can be used to measure terrain features such as elevation (freely available from the Shuttle Radar Topography Mission (SRTM) and other sources). Images of Earth at night can be used as proxies for a region's access to electricity (De Juan and Bank 2015) or for more general economic development (Min 2009). Such data are available from the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) provided by the US National Centers for Environmental Information (NCEI). Finally, data on precipitation and on soil properties can provide information about access to renewable natural resources or the agricultural potential of a given territory (see for example the data sets provided by the GeoNetwork of the United Nations Food and Agriculture Organization (FAO)). A particular advantage of these data is that the probability of any intentional manipulation is very low. Many authoritarian regimes have been particularly reluctant to provide meaningful socio-economic information; however, remote sensing information is not within reach of these states.

As regards the more general challenges of measurement and operationalization beyond the problems related to the lack of available data, subnational research designs may actually offer certain advantages when compared with cross-country analyses. Geographical disaggregation permits the researcher to use indicators that are closer to the local and individual levels. For instance, subnational analyses have relied on information concerning households' durable assets per administrative unit as an indicator of subnational poverty and rates of economic inequality (e.g. Hegre *et al.* 2009; Østby *et al.* 2011; Fjelde and Østby 2012). Such operationalization can more appropriately capture concepts and hypotheses related to the role of poverty and deprivation in political violence than can nationally aggregated measures such as GDP levels. One possible strategy for dealing with the remaining uncertainties involves availing oneself of a specific opportunity that single-country subnational analysis provides; that is, identifying and using context-specific indicators rather than drawing on more generic catch-all proxies. Case-specific qualitative methods can be used to scrutinize the validity of quantitative indicators and to permit informed judgements as to whether these indicators persuasively reflect the underlying theoretical concepts in the specific case under investigation.

In the case of my study of Darfur, I used information from remote-sensing sources concerning environmental conditions to measure the primary explanatory variables. The Normalized Difference Vegetation Index (NDVI) is generated from satellite imagery. Data stem from the Global Inventory Modeling and Mapping Studies' (GIMMS) 15-day Advanced Very High Resolution Radiometer (AVHRR) and are provided by the Global Land Cover Facility at the University of Maryland (see Tucker *et al.* 2004). The NDVI exploits the fact that

healthy vegetation strongly reflects the near-infrared part of the electromagnetic spectrum and absorbs wavelengths of visible red light. As vegetation cover changes, so do these reflectance properties. Thus, satellite imagery can be used to estimate the relative density and health of vegetation, as well as changes over time. Higher NDVI values indicate greater vegetation density and health. I aggregated NDVI data by grid cell to generate variables for environmental change prior to the beginning of the War in Sudan in 2003. Previous in-depth studies on Darfur revealed that vegetation in the region I studied is largely determined by the three-month rainy season (from July to September). Rainfall during these three months is around 90 per cent of the total annual rainfall. I therefore focused my measurement on this specific period, using the average NDVI values for these three months. Using QGIS, I calculated per grid cell the mean NDVI value of all pixels included in the respective cell. In addition, I relied on the aforementioned ELVM study for most of the control variables considered in the main models. This approach allowed me to make use of comparatively reliable data on a geographically highly disaggregated level.

Interpreting the results

Even with convincing hypotheses, a careful selection of units of analysis and reliable data sources, statistical correlations may still be misleading. In the case of my own study, I was able to document a robust correlation between long-term environmental change and the occurrence and intensity of violence in line with my theoretical expectations. But how could I be sure that these findings would actually reflect my specific arguments concerning the link between migration from areas in which the availability of natural resources was decreasing to areas in which it was increasing and the ensuing ethnic conflict in these regions? This challenge is not unique to geospatial subnational analysis – that is, it is a challenge that confronts every purely observational study – but it is of particular concern for this specific methodological approach: the need for auxiliary assumptions to connect causes of rebel activity to the location of the observable violent events introduces additional uncertainty and makes causal claims even more challenging.

A number of technical approaches can be used to further scrutinize statistical correlations. In the absence of a robust experimental design, however, such strategies cannot be employed to reliably dispel any remaining doubts about the actual mechanisms at play. Consequently, it may be more promising to make use of a multi-method design. Such an approach would allow subnational geospatial analyses to contribute their specific added value while their potential weaknesses would be offset by qualitative analyses on the local level. Techniques such as process tracing (see Chapter 8 by Andreas Kruck) can provide important additional layers of evidence in favour of or against the causal hypotheses investigated in statistical analyses (Raleigh *et al.* 2010b). Such a multi-method design could be set up as an actual nested analysis in which the subnational statistical analysis of administrative units or grid cells is combined

with an in-depth investigation of one or more of the units contained in the larger sample (see Lieberman 2005).

In the case of Darfur, I decided to scrutinize the suggested auxiliary hypotheses linking environmental change to militia attacks on villages by making use of a large body of qualitative and descriptive quantitative data obtained from previous ethnographic studies and historical archives (original maps of ethnic settlement patterns provided by the British colonial government), as well as the remote-sensing systems mentioned above, to connect individual elements of the suggested causal chain. Most notably, I showed that (1) environmental change has been highly uneven in Darfur; (2) various ethnic groups have been affected by environmental change to varying degrees; (3) decreasing resource availability led to migration of affected groups into more affluent regions; (4) these movements increased inter-ethnic tensions in areas of high in-migration; and that (5) these tensions motivated militia formation and attacks. Obviously, these step-wise qualitative accounts do not permit strong causal claims, but they do lend support to the argument that these processes link environmental change to observable patterns of militia violence.

Conclusion

Geospatial analyses on the subnational level can shed new light on hypotheses that have been tested in country-level analyses and can make it possible to consider aspects of NSA that have thus far been widely neglected. This approach is best suited to addressing research questions that focus on variables with a high degree of within-country variation and with a strong spatial component. However, when it comes to identifying plausible spatial hypotheses, this specific methodological approach also presents its own specific challenges, the most notable of which are the selection of appropriate units of analysis, operationalization and measurement, and the interpretation of statistical correlations.

To mitigate these challenges and to optimize the specific benefits to be gained through geospatial analysis, the following recommendations should be kept in mind. First, when conducting analyses, researchers should generally think twice before transferring theoretical arguments and empirical findings from the cross-country level to the geospatial level, and vice versa. It may be more promising to use geospatial analysis for particular research questions that are geared more specifically to the actual geographical characteristics of NSA features and activities and to develop specific spatial hypotheses. Second, researchers should evaluate the plausibility of implicit or explicit assumptions that relate observable violence in one subnational region to specific features of that region. The underlying assumptions should be made explicit and justified theoretically and/or empirically. Third, units of analysis should ideally be selected on explicit conceptual and theoretical grounds. Purely data-driven selections of units of analysis should be made explicit. To demonstrate that empirical results are not driven by the researcher's choice of the unit of analysis, additional robustness checks should be conducted using alternative (in terms of size and location) units of

analysis. Fourth, to reduce uncertainty resulting from potentially ambiguous indicators, researchers should invest time in identifying context-specific data sources rather than relying on conventional, readily available indicators of dubious validity. Finally, researchers should consider the potential of nested analyses that combine quantitative and qualitative analyses as a way to ensure that quantitative findings in subnational geospatial analysis are interpreted in a meaningful way.

Note

- 1 The following textbooks provide more general technical introductions to the use of geographical information systems and statistical analyses of spatial data: Isaaks and Srivastava (1989); Cressie (1993); Bossler *et al.* (2002); Heywood *et al.* (2011).

References

- Bossler, John D., John R. Jensen, Robert B. McMaster, and Chris Rizos, eds. 2002. *Manual of Geospatial Science and Technology*. London: Taylor & Francis.
- Bueno de Mesquita, Ethan, C. Christine Fair, Jenna Jordan, Rasul Bakhsh Rais, and Jacob N. Shapiro. 2013. *The BFRS Political Violence in Pakistan Dataset*. Version 7 July 2013. Available from: <https://esoc.princeton.edu/files/bfrs-political-violence-pakistan-dataset> (accessed 15 February 2016).
- Buhaug, Halvard. 2010. 'Dude, Where's My Conflict? LSG, Relative Strength, and the Location of Civil War'. *Conflict Management and Peace Science* 27(2): 107–128.
- Buhaug, Halvard, and Scott Gates. 2002. 'The Geography of Civil War'. *Journal of Peace Research* 39(4): 417–433.
- Buhaug, Halvard, and Jan Ketil Rød. 2006. 'Local Determinants of African Civil Wars, 1970–2001'. *Political Geography* 25(3): 315–335.
- Buhaug, Halvard, Kristian Skrede Gleditsch, Helge Holtermann, Gudrun Østby, and Andreas Forø Tollefsen. 2011. 'It's the Local Economy, Stupid! Geographic Wealth Dispersion and Conflict Outbreak Location'. *Journal of Conflict Resolution* 55(5): 814–840.
- Cederman, Lars-Erik, Halvard Buhaug, and Jan Ketil Rød. 2009. 'Ethno-Nationalist Dyads and Civil War: A GIS-Based Analysis'. *Journal of Conflict Resolution* 53(4): 496–525.
- Costalli, Stefano, and Francesco N. Moro. 2011. 'The Patterns of Ethnic Settlement and Violence: A Local-Level Quantitative Analysis of the Bosnian War'. *Ethnic and Racial Studies* 34(12): 2096–2114.
- Cressie, Noel A.C. 1993. *Statistics for Spatial Data*. New York: Wiley.
- De Juan, Alexander. 2015. 'Long-Term Environmental Change and Geographical Patterns of Violence in Darfur, 2003–2005'. *Political Geography* 45: 22–33.
- De Juan, Alexander, and André Bank. 2015. 'The Ba'athist Blackout? Selective Goods Provision and Political Violence in the Syrian Civil War'. *Journal of Peace Research* 52(1): 91–104.
- De Luca, Giacomo, Jean-François Maystadt, Petros G. Sekeris, and John Ulimwengu. 2012. *Mineral Resources and Conflicts in DRC: A Case of Ecological Fallacy*. CRED Working Paper No. 1207. Namur: Center for Research in the Economics of Development University of Namur.

- Eck, Kristine. 2012. 'In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events Datasets'. *Cooperation and Conflict* 47(1): 124–141.
- Fjelde, Hanne, and Gudrun Østby. 2012. 'Economic Inequality and Inter-Group Conflicts in Africa'. Paper prepared for presentation at the Democracy as Idea and Practice Conference, Oslo, 12–13 January 2012.
- Gleditsch, Kristian S., and Nils B. Weidmann. 2012. 'Richardson in the Information Age: Geographic Information Systems and Spatial Data in International Studies'. *Annual Review of Political Science* 15: 461–481.
- Gurr, Ted R. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
- Hegre, Håvard, Gudrun Østby, and Clionadh Raleigh. 2009. 'Poverty and Civil War Events: A Disaggregated Study of Liberia'. *Journal of Conflict Resolution* 53(4): 598–623.
- Heywood, D. Ian, Sarah Cornelius, and Steve Carver. 2011. *An Introduction to Geographical Information Systems*. 4th edn. Harlow: Prentice Hall.
- International Organization of Migration and ProAct Network. 2010. *Environmental and Livelihoods Vulnerability Mapping (ELVM) in North and South Darfur, Sudan: Final Report*. Available from: www.preventionweb.net/files/15579_proactreport.pdf (accessed 15 February 2016).
- Isaaks, Edward H., and R. Mohan Srivastava. 1989. *Applied Geostatistics*. New York: Oxford University Press.
- Kalyvas, Stathis N. 2003. 'The Ontology of "Political Violence": Action and Identity in Civil Wars'. *Perspectives on Politics* 1(3): 475–494.
- Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.
- Kalyvas, Stathis N. 2008. 'Promises and Pitfalls of an Emerging Research Program: The Microdynamics of Civil War', pp. 397–420 in *Order, Conflict, and Violence*, S.N. Kalyvas, I. Shapiro, and T. Masoud, eds. Cambridge: Cambridge University Press.
- King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton: Princeton University Press.
- Lee, Barrett A., Sean F. Reardon, Glenn Firebaugh, Chad R. Farrell, Stephen A. Matthews, and David O'Sullivan. 2008. 'Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographic Scales'. *American Sociological Review* 73(5): 766–791.
- Lieberman, Evan S. 2005. 'Nested Analysis as a Mixed-Method Strategy for Comparative Research'. *American Political Science Review* 99(3): 435–452.
- Min, Brian. 2009. 'Distributing Power: Public Service Provision to the Poor in India'. Paper presented at the 2009 American Political Science Association Meeting in Toronto, Ontario.
- Montalvo, Jose G., and Marta Reynal-Querol. 2007. *Ethnic Polarization and the Duration of Civil Wars*. Post-Conflict Transitions Working Paper No. 6. Washington, DC: World Bank.
- Oberwittler, Dietrich, and Per-Olof H. Wikström. 2009. 'Why Small Is Better: Advancing the Study of the Role of Behavioral Contexts in Crime Causation', pp. 35–58 in *Putting Crime in Its Place: Units of Analysis in Geographic Criminology*, D. Weisburd, W. Bernasco, and G.J.N. Bruinsma, eds. New York: Springer.
- O'Loughlin, John. 2007. 'Spatial Analysis in Political Geography', pp. 30–46 in *A Companion to Political Geography*, J. Agnew, K. Mitchell, and G. Toal, eds. Malden: Blackwell.
- Openshaw, Stan. 1996. 'Developing GIS-Relevant Zone-Based Spatial Analysis Methods', pp. 55–74 in *Spatial Analysis: Modelling in a GIS Environment*, P. Longley and M. Batty, eds. Cambridge, UK: Wiley.

- Østby, Gudrun, Henrik Urdal, Mohammad Zulfan Tadjoeeddin, S. Mansoob Murshed, and Håvard Strand. 2011. 'Population Pressure, Horizontal Inequality and Political Violence: A Disaggregated Study of Indonesian Provinces, 1990–2003'. *Journal of Development Studies* 47(3): 377–398.
- Raleigh, Clionadh. 2011. 'Scales of Conflict Research'. *International Interactions* 37(4): 469–481.
- Raleigh, Clionadh. 2012. 'Violence against Civilians: A Disaggregated Analysis'. *International Interactions* 38(4): 462–481.
- Raleigh, Clionadh, and Håvard Hegre. 2009. 'Population Size, Concentration, and Civil War: A Geographically Disaggregated Analysis'. *Political Geography* 28(4): 224–238.
- Raleigh, Clionadh, and Dominic Kniveton. 2012. 'Come Rain or Shine: An Analysis of Conflict and Climate Variability in East Africa'. *Journal of Peace Research* 49(1): 51–64.
- Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joakim Karlsen. 2010a. 'Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature'. *Journal of Peace Research* 47(5): 651–660.
- Raleigh, Clionadh, Frank D.W. Witmer, and John O'Loughlin. 2010b. 'A Review and Assessment of Spatial Analysis and Conflict: The Geography of War', pp. 6534–6553 in *The International Studies Encyclopedia, Vol. X*, R.A. Denmark, ed. Oxford, UK: Wiley-Blackwell.
- Reuveny, Rafael. 2007. 'Climate Change–Induced Migration and Violent Conflict'. *Political Geography* 26(6): 656–673.
- Robinson, William S. 1950. 'Ecological Correlations and the Behavior of Individuals'. *American Sociological Review* 15(3): 351–357.
- Rød, Jan Ketil, and Halvard Buhaug. 2007. 'Civil Wars: Prospects and Problems with the Use of Local Indicators'. Unpublished paper.
- Rustad, Siri C.A., Halvard Buhaug, Åshild Falch, and Scott Gates. 2011. 'All Conflict Is Local: Modeling Sub-National Variation in Civil Conflict Risk'. *Conflict Management and Peace Science* 28(1): 15–40.
- Salmon, Jago. 2007. *A Paramilitary Revolution: The Popular Defence Forces*. HSBA Working Paper No. 10. Geneva: Small Arms Survey, Graduate Institute of International and Development Studies.
- Schutte, Sebastian, and Nils B. Weidmann. 2011. 'Diffusion Patterns of Violence in Civil Wars'. *Political Geography* 30(3): 143–152.
- Stephennne, Nathalie, Clementine Burnley, and Daniele Ehrlich. 2009. 'Analyzing Spatial Drivers in Quantitative Conflict Studies: The Potential and Challenges of Geographic Information Systems'. *International Studies Review* 11(3): 502–522.
- Sundberg, Ralph, and Erik Melander. 2013. 'Introducing the UCDP Georeferenced Event Dataset'. *Journal of Peace Research* 50(4): 523–532.
- Tucker, Compton J., Jorge E. Pinzon, and Molly E. Brown. 2004. *Global Inventory Modeling and Mapping Studies, NA94apr15b.n11-VIg, 2.0*. College Park: Global Land Cover Facility, University of Maryland.
- Zhukov, Yuri M. 2012. 'Roads and the Diffusion of Insurgent Violence: The Logistics of Conflict in Russia's North Caucasus'. *Political Geography* 31(3): 144–156.

11 Discussion chapter

Shadow boxing in Plato's cave: assessing causal claims

Bertjan Verbeek

Introduction

The purpose of this contribution is to assess how several of the methodological approaches to investigating the role of non-state actors in international security, introduced and applied in this volume, contribute to the possibility of stating plausible causal claims. To that effect (no pun intended) this chapter dissects three contributions that specifically address the issue of causal claims: comparative case study methodology (Chapter 8 by Andreas Kruck on the privatization of security); qualitative case analysis (Chapter 9 by Patrick Mello on armed conflict); and geospatial analysis (Chapter 10 by Alexander De Juan on the emergence of violent conflict). I will organize my own claims around several questions to be put to each individual chapter: how to select a suitable method for answering a particular research question? How to select cases that fit research design and method(s)? What type of knowledge is served, or hindered, by our choice of method(s)? What can be said to have been explained? I will discuss the three chapters in this part of the volume while keeping these questions in mind. As will be shown, there is not always *the* perfect solution: researchers often have to make decisions that involve trade-offs.

One preliminary observation is in order, however: by assessing whether and how specific methodologies allow for causal claims it is presupposed that, at least in principle, causality exists, and that, to different degrees and possibly only indirectly, evidence exists that allows us to agree on the plausibility of causality to have occurred (see Guzzini 2016). Within this spectrum many different positions are still tenable and indeed the International Relations (IR) discipline boasts hugely diverse notions of causality. On the one hand, we find those who, sometimes explicitly modelling themselves after the natural sciences, seek for universal laws of international politics – Waltz's claims about balancing in anarchical systems come closest to that notion (Waltz 1979). On the other hand, we find those who reject the possibility of universal laws, but still agree that even unique historical developments can be unpacked as a sequence of events representing a causal chain in which one phenomenon at t_1 helped bring about another at t_2 .

Social Constructivists often find themselves at odds with either position: because universal laws presuppose similar, if not identical, actors and contexts

to be present in the international realm, constructivists argue that actors, their properties, and the contexts in which they find themselves change over time because their social existence (constitution) is subject to social construction. Although this leaves little room for universal laws, for some constructivists this does not render impossible the notion of causality as long as the scholar strives for causal claims about time- and space-bound, comparable phenomena. Hence their emphasis on the importance of so-called scope conditions (e.g. Checkel 2001).

Still, a constructivist would argue, another issue has to be tackled before one can engage in such ‘small t theorizing’: the issue of mutual constitution. Because, in *social* constructivism, meaning emerges out of interaction between two or more actors, such meaning is mutually constituted in a process in which it is difficult to assume a clear linear process of x leading to y . Rather, for a constructivist, y affects x while, at the same moment, x affects y . This hugely complicates the issue of causal claims, because dissecting the mutual constitutive effect still requires the adoption, if only for pragmatic reasons, of some kind of time sequence. Hence, the position of some constructivists that constructivist research requires engaging in so-called bracketing. Bracketing involves the temporary assumption that one side of the mutual interaction can be considered to be constant in order to describe the other side. In this sense, constructivists borrow from an experimental research design (Price and Reus-Smit 1998; for a different perspective see Fierke 2005: 9–16; Klotz and Lynch 2007: 14–15).

Critical theorists would not necessarily discard the possibility of causal claims. They would point to two major issues: first, because to them empirical phenomena are by definition context bound (in time and place) they would hesitate talking in terms of universal laws and strive for limited causal claims for a circumscribed class of phenomena. Second, they would emphasize that the individual researcher herself is part of the social phenomenon she is studying. They thus add a new dimension to causality: everything the researcher does or does not do may have a causal effect on the phenomenon to be studied. This is diametrically opposed to those scholars in our field (such as, for example, neo-realists) who would argue that the investigator is looking in on international reality from the outside and that her presence does not affect the subject under study. A critical theorist would therefore require the scholar to be aware of her position and impact on the matter under investigation.

With these caveats in mind, I will look upon the three chapters on causal claims. My own position is that causality implies a time sequence (even if in the form of constructivist ‘bracketing’) and that because of this we are implicitly or explicitly thinking in terms of an independent variable preceding a dependent variable. In addition, I start from the notion that establishing a causal claim not only requires an observation of a sequence of events, but above all an explanatory mechanism accounting for *why* the sequence has occurred. Such explanatory mechanisms can only be found in theories, also if we are investigating phenomena that only occur in a specific context of time and space.

How to select a suitable method for answering a particular research question?

One often hears the suggestion that one's research question determines one's research method. Yet, that claim is still no shortcut to a pragmatic decision on method. First of all, we need to distinguish between research design and methods in the more narrow sense. Second, we need to distinguish methods to collect data from methods to analyse data. One's research design is related to one's conceptual model which in turn reflects the research question as well the essential elements of the theory one considers plausible for tackling the research question. Theory could be conceived in a narrow sense (assumptions leading to the deduction of testable hypotheses) or in a wider sense (a set of related concepts that delineate the empirical phenomena we wish to study), but it always implies some kind of reduction of complexity as well as applying an interpretive filter that allows us to make sense of the world. E.g., realists do this when they present world politics as an anarchic system of states (despite the presence of other actors and forms of governance) on the grounds that only states have the capacity to organize in such a way that they can survive in such a surrounding.

Andreas Kruck's research (see Chapter 8) seeks to explain the variation in degree and nature of privatization of security in four different countries. Andreas Kruck opts for a comparative case study of France, Germany, the United Kingdom and the United States. Now the research question itself does not preclude other methods: it allows for a large- n study of privatization. His operational definition of privatization would certainly facilitate that. Yet, Andreas Kruck seeks to accomplish more than correlation between variables: he wants to see *whether* the causal mechanisms underlying such a correlation can be unveiled. In order to trace causality, one could, instead, consider opting for a large- n study supplemented with one or more case studies, a strategy sometimes preferred if a large n is available (Landman 2005). Andreas Kruck, however, opted for four case studies in order to establish both the 'comprehensive explanation' and the causal claims. He might have done so for pragmatic reasons because large-scale standardized data for numerous countries that would be necessary for a large- n study (supplemented by case studies) were not available as ready-made data sets and might be hard to gather by the individual researcher. He calls this 'zooming in' and 'linking up'. This is a plausible strategy for situations in which a large n is absent and a comparative case study seems warranted. At any rate, the broader lesson is that the number of (potentially) available n is an important determinant for the choice of an adequate method. The number of available cases may in turn depend on a whole range of factors including the real world frequency of a phenomenon, availability of (reliable) data and the researcher's time and resources for gathering new data. So researchers might want to ask themselves: would my research question allow for large n analyses, or does it suggest a comparative case study? Which additional considerations regarding data collection and data analysis would this demand from my research?

Patrick Mello (see Chapter 9) is looking for an explanation of the choices by 30 democracies of whether or not to participate in the 2003 Iraq War. Patrick Mello suggests Qualitative Comparative Analysis (QCA) as an appropriate method to tackle this specific question. Although his research involves few – if any – non-state actors, his empirics indicate that QCA can be a viable avenue for specific forms of research into non-state actors. QCA is a method of analysis that enables us to establish whether individual variables or specific combinations of variables constitute necessary or sufficient conditions for a specific outcome to occur by formally attaching weights to qualitative variables and counting the actual occurrence of all possible logical combinations of independent variables. Now QCA offers particular advantages to comparative case studies: when n is too small to engage in probability tests, yet the number of cases is substantive, QCA offers a method to make a statement about causal links between independent and dependent variables (or conditions and outcomes, as researchers using QCA prefer to call them) – importantly, it is a method that facilitates replication. The additional advantage is that QCA allows for making plausible statements about there being various causal pathways to the same outcome (equifinality). Furthermore, QCA permits identifying clear outlier cases, which could then be investigated in depth through single case studies. Hence, it might be useful for a lot of studies of non-state actors in international security, if researchers asked themselves: does my research allow for the possibility of different causal pathways leading to the observable outcome? Would in my research design QCA be useful to depict various such pathways?

Alexander De Juan (see Chapter 10) seeks to investigate the relationship between reduced availability of natural resources and the advent and intensity of violent conflict. His analysis involves data pertaining to the situation in the Sudanese region of Darfur. Interestingly, Alexander De Juan's research question springs from an empirical observation that geographical patterns of violent conflict do not coincide with geographical patterns of deteriorating supply of vital natural resources. This contradicts the common-sense idea that people would be prepared to use violence over extreme lack of access to (dwindling) resources. The observation induces him to formulate a theory that basically states that dwindling resources force people to migrate to areas of plenty and that this migration causes tension between old and new users of these resources. This observation allows for a conceptual model accounting for these two movements (first, migration patterns; second, use of violence) and for readily testable hypotheses, which could be approached both with qualitative and with quantitative methods. Alexander De Juan's approach thus highlights the importance of formulating testable hypotheses. This points to another set of questions relevant for causal research: does my research design allow me to formulate straightforward expectations? Which type of methods would allow me to put these to the test?

Data collection and units of analysis

In terms of data collection, it requires a decision on the unit of analysis: the case to be selected is supposedly part of a class of phenomena about which one's theory is expected to provide a meaningful interpretation. For Andreas Kruck (see Chapter 8) this seems to be all 'Western "strong" and democratic' countries that have to some extent privatized security. From that group he selects four. Although quite a considerable number for a comparative case study, one could debate whether more may have been better. However, researchers often have to limit their analysis for pragmatic reasons (how much time do I have for my research? What data are available?). Andreas Kruck offers three explanatory theories (functional, political and ideational). In particular, from the perspective of 'covariation analysis' – one of the three case study techniques Andreas Kruck discusses – one might argue that establishing a hint of causation would require holding constant two of three 'theories' each time one 'theory' is 'put to the test'. Four case studies would thus not yet yield sufficient insight into the relative strength of the three theoretical approaches. The four case studies do however yield a sizeable number of empirical observations (beyond IVs and DV) which – through 'congruence analysis', the second case study technique he mentions – can be compared with the predictions of the three theories. Moreover, Andreas Kruck's case selection would allow for tracing the causal mechanism within each approach, but then one case per approach would suffice. The question then is which case comes closest to representing each theoretical approach: the expectations formulated by Andreas Kruck make France the best candidate for the political approach (weakest parliament); the UK for the ideational approach (as the most neoliberal system) and possibly the USA for the functional angle (being the system most geared toward cost saving). Indeed, one would thus engage, as did Andreas Kruck, with structured-focused comparison as a research method (George 1979a): 'focused' implies a plausible link with the theory that informs the research (here: functional, political and ideational accounts of privatization). 'Structured' implies that the selected cases still have to be argued in terms of comparability: certainly, in Andreas Kruck's chapter, France, the UK and USA have, or seek, a leading role in world politics and thus would have a comparable 'interest' in an effective security apparatus. This would be less plausible for Germany. Ultimately, however, achieving both goals of comprehensive explanation and zooming in may profit from additional cases. This specific study thus raises general questions for all kinds of causal research projects: is it clear from my research design how many different theoretical claims are under investigation? Is it clear at what level of analysis my theories seek to account for phenomena? How large is the universe of cases related to these theories? When juxtaposing various theories, how to ensure that my selected cases are, on the one hand, comparable, yet differing in terms of the independent variables my theories identify?

Alexander De Juan (Chapter 10) would like to be able to make claims about all non-state actors engaged in violent conflicts. His theory is about situations in

which natural resources are dwindling. Of course, what constitutes salient resources differs across countries and even within countries and what counts as a salient resource may partly be the product of social construction. Nevertheless, Alexander De Juan seeks to engage with the literature that focuses on natural resources in a traditional sense (related to land, water and air) often in the context of climate change (Buhaug *et al.* 2010). Importantly, Alexander De Juan persuasively argues that resource deterioration, migration and conflict often occur at the subnational level. He rightfully pleads for the collection and analysis of data at the subnational, i.e. regional, level. He proposes the use of geospatial analysis as a method to collect and interpret such regional data, in his case from Darfur in Sudan. This disaggregation is certainly welcome, but if one assumes that actors or structures at the national level affect the situation and choices at the regional level, this should be included in the conceptual model and one should seek to hold constant the impact of the national level. Alexander De Juan elaborately discusses the pros and cons of different approaches to measure phenomena at the regional level and persuasively argues to use specifically sized grids as the optimal unit of observation (rather than administrative units, etc.). Rather than opting for single size grids, he proposes two different-sized circles around villages. Generally, choosing the ‘right’ level of analysis is a crucial but often tricky task, not only but especially for researchers of non-state actors in international security who have to ask themselves: to what extent does my research design give a lead to the level of analysis at which my theory offers an explanation? What does this imply for the actual unit of observation I select in my empirical research? In selecting cases can I control for the impact of variables at different levels of analysis?

Patrick Mello’s (Chapter 9) research question makes the state the unit of analysis. Different theories point to variables at different levels (e.g. realism to military power; liberalism to parliamentary veto power). The advantage of QCA is that combinations of variables at different levels can point to causal pathways, thus allowing for the identification of the complementarity of theories (which may not be the same as the desired integration of theories). Yet the specific research question comes at a price: Patrick Mello’s research question lumps together the decision to participate in the 2003 intervention in Iraq and the decision to contribute (militarily) to the post-war reconstruction. This obscures the position of countries such as Italy and the Netherlands, which opted for political support for the war and, later, for military support for reconstructing Iraq. In this case this would affect the so-called truth table (Table 9.5) and the conclusions regarding necessary and sufficient conditions. It does not jeopardize the valuable contribution of QCA *per se*, but it once more underlines the vital interest of linking precise research questions to theory before the use of specific methods can be assessed. Researchers should therefore always ask themselves: is my research question sufficiently specific about the phenomenon to be accounted for? Do my theories speak to precisely this phenomenon?

The objective of assessing the causal mechanism implies a data analysis method that would allow one to do just that. Andreas Kruck has chosen *process*

tracing to accomplish that aim (see George 1979b; Checkel 2008). His strategy of assessing whether covariation over time or across countries takes place allows him to draw plausible conclusions regarding causal links. Andreas Kruck complies with the prescription that process tracing requires us to pose the same questions to comparable cases and to standardize the data collection of all cases. In doing so, he avoids the problems of thick description, which may give valuable qualified and more complex information, but which entails the risk of losing sight of arguing that causal patterns can or cannot be observed.

The unravelling of causal mechanisms is, however, a challenging endeavour for researchers. The conceptual model that was chosen by Alexander De Juan, for example, cannot yet fully account for the causal chain: we know from the mobilization literature that the step from frustration to aggression requires the incorporation of processes of organization and mobilization (see Tilly 1978); we need to establish whether migrants moved because of the threat rather than for other reasons; we need to know whether variation in conflict in migration areas can be linked to differences in organization and mobilization. Still, in principle, such issues could be measured by taking the geospatial angle, provided data at the grid level are available.

The broader lesson from these examples is that (the growing number of) researchers who are interested in thoroughly studying causal mechanisms will have to address a number of crucial questions: does my research design specify the causal mechanism(s) that may underlie the theoretical approaches to answering my research question? Does my chosen method of analysing data, and the data itself, allow me to draw conclusions about the presence of the identified causal mechanisms? Are my selected cases sufficiently comparable that inferences about causal mechanism can be made?

In Patrick Mello's QCA approach, causality is effectively defined as identifying necessary and sufficient conditions. This renders important insights into his scientific problem: he identifies two pathways to participation in the 2003 Iraq War: the presence of a right-wing executive combined with absence of constitutional restrictions, and the absence of military strength combined with the absence of constitutional restrictions combined with the presence of parliamentary veto rights. Despite these results it remains unclear how these pathways represent the causal mechanisms belonging to the four theoretical approaches Mello sets out to test and, at some point, integrate: institutionalism, realism, liberalism and constructivism. Undoubtedly, this is addressed in the research from which he draws (Mello 2014), but the exact link with theory could have been developed more. Such a shortcoming is a frequent lacuna; maybe because linking empirics to theory is a very demanding task. If this link is underdeveloped, however, it may render it difficult to fully assess the explanatory power of the theories used (or of alternative ones). In the present QCA-chapter this makes it more difficult to assess under what circumstances QCA needs to be supplemented with process tracing in order to detect the causal mechanisms that can be traced back to theory (see George and Bennett 2005: 163). Again, this underlines the usefulness of a research design that

specifies the causal mechanism(s) that may underlie the theoretical approaches to answering one's research question.

What kind of knowledge is generated?

Andreas Kruck has succeeded in plausibly arguing that causal patterns have been observed – the goals of zooming in. Yet, it in order to meet the wider objective of a comprehensive explanation it might be useful to specify more precisely the wider class of phenomena and to place the theories that guide the case selection in the context of larger debates in International Relations about what behaviour to expect from states towards privatization of security, e.g. whether sovereign states can be expected to privatize security in the first place. It is interesting that new puzzles and blind spots are identified but it remains unclear whether they speak to the original research question. However, as said before, researchers often have to limit their analysis due to practical reasons, which can lead to trade-offs. From Alexander De Juan's analysis we learn that geospatial analysis helps us falsify the direct link between resources depletion and violent conflict and establishes the plausibility of an indirect link via migration behaviour to resource-plenty areas. Geospatial analysis enables us to expand the conceptual model (to account for causal mechanisms) to the extent that data at the grid level are available. Generalization of knowledge would nevertheless require controlling for influences at the national (and possibly transnational) level.

Patrick Mello's research underlines the possibility of making causal claims even when the number of cases is too low for quantitative analysis. Because QCA forces the researcher to be extremely explicit about operational definitions and the attaching of weights to variables it facilitates generalization. Although he does not discuss whether his empirical case (states' decisions regarding the 2003 Iraq War) is representative for all armed conflict, it is safe to say that the two causal pathways are plausibly valid for democracies' decisions to go to war in the post-Cold War era. The additional advantage of QCA in identifying outlier cases will help us to conduct supplementary in-depth case studies that trace the conditions that trigger deviant choices. These reflections on the three specific studies invite a final set of questions that will be relevant for many, if not all research projects seeking to make causal claims: is it clear in my research which class of phenomena I seek to theorize? Is it clear whether I consider them time- and space-bound and how that may limit the size of the phenomenon I seek to account for? Is it clear how this relates to the larger debates in IR theory?

Conclusion

First and foremost, the choice of method(s) follows one's conceptual model, which itself is based upon a theory-driven research question. In that sense, each method applied in this section of the volume may be a wise strategy depending on the problem at hand. The chapters in this volume make clear, however, that research question and conceptual model merit further explicit deliberation and

justification. This is necessary to be able to determine the nature of the knowledge generated (e.g. in terms of generalizability) and to assess whether the chosen methods of data collection and data analysis actually make forms of causal claims possible. Although all three methods (Structured Focused Comparison; Geospatial Analysis; QCA) seriously offer a tool to make plausible claims about causality, the chapters show that the former two, but possibly also the latter, may need additional tools (such as process tracing) to reveal underlying causal mechanisms. In order to be able to do so, we need to be more explicit about the causal mechanism that our theories suggest. All in all then, although we may have come closer to making inferences about causality, we may still be in Plato's cave observing shadows.

References

- Buhaug, Halvard, Nils Petter Gleditsch, and Ole Magnus Theisen. 2010. 'Implications of climate change for armed conflict', pp. 75–101 in *Social Dimensions of Climate Change: Equity and Vulnerability in a Warming World*, R. Mearns and A. Norton, eds. Washington: The World Bank.
- Checkel, Jeffrey T. 2001. 'Why Comply? Social Learning and European Identity Change'. *International Organization* (55)3: 553–588.
- Checkel, Jeffrey T. 2008. 'Process Tracing', pp. 114–127 in *Qualitative Methods in International Relations. A Pluralist Guide*, A. Klotz and D. Prakash, eds. Houndmills: Palgrave.
- Fierke, Karin M. 2005. *Diplomatic Interventions. Conflict and Change in a Globalizing World*. Houndmills: Palgrave.
- George, Alexander L. 1979a. 'Case Studies and Theory Development: the Method of Structured Focused Comparison', pp. 43–69 in *Diplomacy. New Approaches to History, Theory and Policy*, P.G. Lauren, ed. New York: The Free Press.
- George, Alexander L. 1979b. 'The Causal Nexus between Cognitive Beliefs and Decision Making Behavior: The "Operational Code" Construct', pp. 95–124 in *Psychological Models in International Politics*, L.S. Falkowski, ed. Westview: Boulder.
- George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Guzzini, Stefano. 2016. 'Power and Cause'. *Journal of International Relations and Development*. doi:10.1057/s41268-016-0002-z.
- Klotz, Audie, and Cecilia Lynch. 2007. *Strategies for Research in Constructivist International Relations*. Armonk: M.E. Sharpe.
- Landman, Todd. 2003. *Issues and Methods in Comparative Politics: An Introduction*. London: Routledge.
- Mello, Patrick A. 2014. *Democratic Participation in Armed Conflict: Military Involvement in Kosovo, Afghanistan, and Iraq*. Houndmills: Palgrave Macmillan.
- Price, Richard, and Christian Reus-Smit. 1998. 'Dangerous Liaisons? Critical International Theory and Constructivism'. *European Journal of International Relations* (4)3: 259–294.
- Tilly, Charles. 1978. *From Mobilization to Revolution*. New York: McGraw-Hill College.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. Reading: Addison Wesley.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Part III

Doing fieldwork



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

12 An ethnographic approach to non-state security

Participant observation among private security officers

Tessa Diphooorn

Introduction

Participant observation has largely defined how anthropologists do research (Agar 1996; DeWalt and DeWalt 2002). In this chapter, I will discuss how I used this method – participant observation – in my research on analysing the daily practices of private security companies and private security officers and their interactions with other security actors throughout the past eight years. Although I mention several disadvantages and obstacles, my main aim is to highlight the numerous advantages of participant observation and the incredibly rich and in-depth empirical data it can yield.

Many anthropologists have used participant observation to analyse the social world of non-state security actors across the globe, such as gangs (Rodgers 2006; Jensen 2008) and vigilante organizations (Goldstein 2012; Pratten and Sen 2007; Buur 2006). Yet ethnographic accounts of commercial private security companies – that are primarily based on participant observation – are scarce (Diphooorn 2016; Konopinski 2014; Rigakos 2002). In fact, when I tell people that I study private security, many question how this is related to anthropology. It seems that the growing and expanding private security industry has failed to catch the attention of anthropologists and has primarily been contained within the fields of criminology and political science.

As noted by Daniel Goldstein (2010: 487), ‘anthropology has not developed a critical comparative ethnography of security and its contemporary problematics’. Furthermore, there is a reigning assumption that participant observation among private security companies inherently places researchers in a range of ethical, emotional, and moral conundrums. Although I do not refute this claim, I also argue that participant observation has numerous benefits, such as gaining access into the daily practices of private security officers, uncovering a bottom-up perspective to security, and understanding how private security officers actually experience their line of work. In fact, I argue that when applied effectively, participant observation provides insights that other methods cannot, such as discovering both what people say and what they do, and the potential differences between the two.

In this chapter, I will discuss both the advantages and disadvantages of using participant observation when studying private security companies and I will

draw from my own field experiences from two different projects on private security companies that took me to South Africa, Kenya, Jamaica and Israel. In the next section, I will briefly outline the two research projects. In the third and largest section, I will examine participant observation as a method and analyse the seven main elements that define it: participation, open interviews, observation, field notes, longitudinal research, establishing rapport, and reflexivity. In this discussion, I will present some of the benefits and obstacles I faced and how I dealt with these issues. In the fourth section, I bring these seven elements together and emphasize that participant observation provides insight into what people say *and* do. This chapter ends with some concluding remarks on the use of participant observation when studying non-state security actors.

Research projects and objectives

This chapter is based on experiences from two different research projects. The first is my PhD research about the armed response sector in Durban, South Africa. The armed response sector is a specific type of private security that comprises armed security officers who patrol communities in vehicles and react and/or respond to triggers such as alarms and panic buttons that are installed on clients' premises. Durban was selected as the research site, because it is a large urban centre, experiences high rates of violent crime, has great demographic diversity, and is less-researched than other South African cities, such as Cape Town and Johannesburg (Diphoorn 2016).

I commenced the research with a rather broad research objective, namely to uncover the daily policing practices of armed response officers and their niche within the larger security spectrum. Employing a pluralized perspective on policing, I was not only interested in finding out who the men are that perform this work and how the companies operate, but also how they interact with other actors, such as clients and police officers. My research approach was thus rather exploratory: it was clear what I was looking at, but it was not clear what I was going to find. This research goal virtually necessitated participant observation as a central research method, as other research methods would not allow me to achieve this objective. My ethnographic approach contrasted with most studies of private security officers, which derive from the field of criminology and employ a quantitative approach to present the 'typical security officer' in a particular space and time (Manzo 2004; Wakefield 2003).

Between 2007 and 2010, I spent a total of 20 months in Durban, spread across three periods: October to December 2007, June 2008 to May 2009, and April 2010 to September 2010. Although I frequented several companies and interviewed the owners of more than 20, I focused on four companies for in-depth analysis. These firms were selected to reflect the diversity of the industry. The first is an internationally owned company that operates globally. The second is a Durban-based company that operates solely in the Durban Metropolitan Area (DMA). The third and fourth are community-based companies: one operates in an affluent, predominantly white area while the other operates in a former Indian

township on the outskirts of the city. Although I used several qualitative methods, the bulk of the data collected for this project derives from participant observation. In addition, I also conducted structured interviews, group interviews with other types of security officers, recorded the life histories of several individuals, and conducted some degree of secondary data analysis, such as employee contracts and minutes of meetings. The other methods were used as a means of triangulation and to support participant observation. For example, when I attended operational meetings organized by the state police, I used the minutes written by someone else and compared them with my own observations.

After my PhD research in South Africa, I started my post-doctoral research within a larger comparative project entitled 'public-private security assemblages', which focuses on the interactions between different security providers and how these impact citizenship in five different cities: Nairobi (Kenya), Recife (Brazil), Jerusalem (Israel–Palestine), Kingston (Jamaica), and Miami (USA).¹ In addition to four PhD researchers, another post-doctoral researcher, and the principal investigator, my role was to conduct a comparative analysis of the composition, regulation and operation of these security assemblages in Nairobi (Kenya), Jerusalem (Israel–Palestine), and Kingston (Jamaica). In contrast to my PhD research, participant observation was not my main research method, and I predominantly used interviews, life histories, and secondary data analysis. The main reason for this concerns the research question, which comprises three parts. The first two parts, namely the 'composition' and 'regulation' of the security assemblages, do not inherently require participant observation, but can be gathered through interviews and data analysis. It is only the third component of the research question, namely the operation of the security assemblage, that requires participant observation. Thus, while in South Africa, participant observation was the key method that was supplemented by others, in this project, it was the other way around: participant observation complemented the other methods. The centrality of participant observation thus differed in each project, further highlighting the versatility of this method.

Participant observation

Although many anthropologists have carried out forms of participant observation before him, Bronislaw Malinowski is regarded as the founder of this method, as he was the first to explicitly outline it as a technique during his fieldwork on the Trobriand Islands in the 1920s. Bronislaw Malinowski's emphasis on everyday interactions and observations and the systematic noting down of this is what differentiated his approach from others (DeWalt and DeWalt 2002: 3).

I employ Kathleen DeWalt and Billie DeWalt's (2002: 1) conceptualization of participant observation, defined as

a method in which a researcher takes part in the daily activities, rituals, interactions, and events of a group of people as one of the means of learning the explicit and tacit aspects of their life routines and their culture.

‘Explicit’ refers to things that people can articulate themselves, such as how they feel and think about something, and ‘tacit’ refers to things that we are not directly aware of, but that steer how we act. The most commonly used examples to explain the latter are the appropriate amount of space between people when standing in a queue or how a wink is interpreted. These are interpretations that we do not think about, but are elements that we simply know and enact.

Participant observation is essentially an oxymoron, as participating and observing are opposing activities. Observation is to detach oneself from the research setting and look at what is happening. Participation, on the other hand, entails the opposite, as one is supposed to immerse oneself in a situation. Building on Spradley’s (1980) typology of participation, Kathleen DeWalt and Billie DeWalt (2002: 19–22) discuss a continuum, including nonparticipation, passive participation, moderate participation, active participation, and complete participation. In my research, *nonparticipation* did not exist – I was always there, and thus, always participating. Furthermore, full participation, complete immersion, or ‘going native’ was also not the case, as I always remained an outsider. For example, when accompanying armed response officers on their shifts, I participated in most of their activities, but not all of them, such as driving the company vehicles and arresting suspects. I therefore did things differently than my research participants and was not fully immersed.

Elsewhere (Diphooorn 2013), I discuss three different modes of participation, namely active participation, reluctant participation and passive participation, during my research in South Africa. The first mode is *active participation*, and I discuss how I experienced giving mouth-to-mouth resuscitation to the family member of an informant. It is an incident where I actively and willingly stepped into my research setting by moving beyond the position of the researcher and concordant to my research population. The second is *reluctant participation*, a mode I often encountered during violent incidents, in which I participated rather unwillingly, less actively, and differently than my informants. The third is *passive participation*, where I deliberately tried to inhabit the role of observer.² These three modes of participation do not refer to different roles or personal identifiers, but they examine the actions I undertook in the field during particular (violent) encounters. It is not a fixed classification, but functions more as a heuristic device to analyse our position, as participants, in relation to other research participants.

There is much debate about what exactly participant observation is, particularly with regards to anthropology as a discipline. Participant observation is often regarded as the means in which anthropologists distinguish themselves from other social sciences. Yet numerous scholars, such as Tim Ingold (2008) and Martin Forsey (2010), highlight that participant observation should not be equated with ethnographic fieldwork: it is a method that largely defines ethnographic fieldwork, but it is one method among many, and almost all anthropologists combine this method with others. Furthermore, the method is increasingly being used in other disciplines, such as criminology and sociology. In policing studies, for example, participant observation is increasingly encouraged and

common (Marks 2004; Punch 1986). In fact, some evocative ethnographies on private security, such as George Rigakos's (2002) excellent account of the Canadian security company Intelligarde, are not conducted by anthropologists, and other scholars studying private military companies are also increasingly using participant observation (Chisholm 2014; Higate 2011). Participant observation is thus not confined to anthropology and can also be used in other disciplines, such as political science.

In the following sections, I will outline how I define participant observation as a larger tool-kit that primarily consists of the following seven elements: (1) actively participating in daily activities and 'hanging out', (2) using and regarding everyday conversations as a form of interviewing, (3) continuously observing one's surrounding, (4) systematically recording the observations made, (5) living in a particular area for a substantial amount of time, (6) establishing rapport, and (7) being reflexive. In the following subsections, I will discuss each element and how I implemented this in my fieldwork.

Participation

The first element concerns participating and 'hanging out'. This entails attending as many activities as possible. In the past, anthropologists often studied entire 'cultures' by conducting research on particular tribes and groups of people. Well-known accounts are Bronislaw Malinowski's work on the Trobriand Islands (1922) and Edward Evans-Pritchard's (1940) account of the Nuer in South Sudan. In these accounts, participant observation entailed participating in all of the events of the people one studied. These studies often occurred in defined geographical areas, where 'the field' involved literally stepping outside one's doorstep. This is increasingly not the case, as my research portrays. Participant observation is more often used to analyse particular aspects of social life, such as security. Furthermore, there is a much more expansive understanding of what 'the field' is (Gupta and Ferguson 1997), where forms of multi-sited ethnography (Marcus 1995), nonlocal ethnography (Feldman 2011), and anthropology at home (Jackson 1987) are increasingly common.

In my research, participation entailed taking part in security-related events, specifically those associated with the armed response sector. This primarily meant accompanying armed response officers during their 12-hour shifts. In total, I spent approximately 750 hours with armed response officers in their vehicles – this amount only includes the actual time spent on their shifts. When I went out on duty with the armed response officers, I wore a bulletproof vest on which I attached my university business card to clearly show that I was a researcher and not an employee of the company. In addition to the 750 hours spent on the road, I regularly hung out at the company offices. This involved finding a seat in the office, looking at what happened around me, and striking up conversations with people walking by. This was slightly awkward at first, but eventually my presence became normal. I also participated in the control room, the technical hub of a company, accompanied technicians and sales reps on their

rounds, completed several private security training courses (including a firearm training course), worked as a car guard for a few days, attended various work-related settings, such as disciplinary hearings and job interviews, tagged along with private investigators and bodyguards, spent approximately 100 hours with citizens on their patrols, accompanied the state police on their shifts for approximately 80 hours, did a polygraph test, and much more. All of these participatory efforts allowed me to gain a richer understanding of private security and crime in South Africa.

Yet participation in a research that involves the use of violence can also place the researcher in dangerous and ethically charged situations and I experienced numerous violent encounters during my fieldwork in South Africa. The troublesome nature of researching violence is not under-analysed in anthropological, criminological and sociological fields (Ghassem-Fachandi 2009; Lee 1995, Lee-Treweek and Linkogle 2000; Punch 1986; Robben and Nordstrom 1995; Rodgers 2006; Sluka 1990; Scheper-Hughes and Bourgois 2004). Researching violence is emotional – for both the informants and the researcher – and it carries additional responsibilities and dilemmas that outweigh those associated with traditional ethnography (Nordstrom and Robben 1995; Ghassem-Fachandi 2009). Existing ethnographies on violence clearly portray the emotional turmoil inherent in such research, and my own work follows suit (Diphoorn 2013, 2016). Feelings of estrangement, frustration, guilt and disgust and the recurrence of nightmares and insomnia were all part of the research experience. I often experienced this as an impediment to the research process and I felt an incessant need to neutralize or subdue certain emotions. One of the means with which I dealt with this process was to repeatedly reflect on my emotions and outline them in my field notes. This not only had a therapeutic effect, but it also allowed me to analyse how my emotions affected my research and changed over time. Another way of coping with these emotions is to maintain frequent contact with colleagues about fieldwork experiences as a means of having an ‘outsider perspective’. Additionally, it is important that supervisors visit their researchers in the field to observe and discuss the emotional process of conducting such research.

Besides emotional distress, research on violence inherently posits a researcher into degrees of judgementalism (Liebling and Stanko 2001; Rodgers 2006). In my view, the use of violence was very often morally wrong, a stance that is palpable in my field notes. Although my primary aim was to understand it, my moral standpoints continuously pierced this process. At the start of my fieldwork, I did not disclose my judgemental opinions about the behaviour of my informants, particularly when they were physically violent. Rather, I focused on asking them about their behaviour in an attempt to understand it; why they acted in such a fashion, what they believed it accomplished, and what they were feeling at the time. Nevertheless, by merely posing these questions and showing the need for clarification, my judgement and dissimilar understandings of morality surfaced. Therefore, when researching violence and security, we need to be aware of the risk we place ourselves in and the ethical and moral predicaments we may face.

When applicable, it is important that ethic committees approve of the research topic and proposed methodology. This is particularly important for students engaging on such endeavours. It is crucial that supervisors outline the necessary precautions, compel students to think about the potential risks they are taking, and how they will deal with certain issues, such as anonymizing the identities of their informants.

Open interviews

The second element that defines participant observation is open interviews. These are interviews that are more like conversations and are not based on a topic list, but are largely guided by the interviewee and take place while participating and observing. For example, when accompanying armed response officers on their shifts in South Africa, we would have endless conversations about their occupation, but also about their personal lives, and other random things, such as music and sports. Similar to Michael Agar (1996), I regard these as a key part of participant observation and they generated the majority of my data, particularly regarding personal issues, such as their relationships with their wives and the difficulties they experienced in 'bringing the work home'. Such conversations also include 'small talk', an essential part of fieldwork and a useful technique when speaking about sensitive issues (Driessen and Jansen 2013). Active listening is also a key component. In fact, Martin Forsey (2010) argues that ethnographic fieldwork is much more about 'participant listening' than participant observation. According to him 'a significant enough portion of ethnographic writing is based more upon what was heard in the field than what is seen there' (Forsey 2010: 563).

Observation

The third important element is to observe one's surroundings. This includes looking at people, such as their bodily behaviour, the physical surroundings, such as the use of space, and the interactions between people. My approach has always been to write down as much as I can remember and incorporate as many details as possible, even if such aspects may seem irrelevant at the time. This is especially pertinent at the start of fieldwork, when it is not clear which details are important. For example, many of my more structured interviews took place in boring office buildings. This may seem like an irrelevant surrounding to describe in comparison to people's homes. Yet even with such places, I describe how I get to that office, what the person is wearing, whether there are particular objects in that office that catch my attention, such as diplomas on the wall or family pictures, etc. Although much of this information may eventually turn out to be unimportant, it can also turn out to be insightful. For example, in Nairobi I interviewed a member of a community policing initiative in his office. While waiting in another room, I saw numerous medals and awards on a shelf that contained his ascribed nickname within that policing initiative.³ By seeing this and noting it down, it made parts of the interview that followed much clearer.

Another technique I habitually use is what Kathleen DeWalt and Billie DeWalt (2002: 69–72) call ‘mapping the scene’, i.e. drawing out the surroundings of events that I attended. Many anthropologists ‘map’ the neighbourhood and area that they research, yet since my research was not focused on a particular neighbourhood, I primarily used this technique when attending certain events. For example, in Durban and Nairobi, I attended operational meetings between police officers and company owners and I always drew out the situation: the size of the room, the table, where everyone was seated, and so forth. Additionally, I would count things, such as the number of police officers versus private security officers, use arrows and colours to note particular interactions between people, and use other measures in order to draw out the different dynamics. These details, particularly when done repeatedly and over time, have provided insight into the dynamics between state police officers and company owners.

Field notes

This technique is also related to the fourth key element of participant observation, namely the recording of field notes. As highlighted by Kathleen DeWalt and Billie DeWalt (2002: 142): ‘observations are not data unless they are recorded in some fashion for further analysis’. We need to remember that participant observation is a method: as humans, we participate in a range of interactions every day, yet it only becomes a method when we systematically note down this information and analyse it for scientific purposes. As mentioned, I contend that field notes should be as detailed as possible, despite how long this may take. Field notes often start from small jotted down notes that consist of keywords. For example, I always carry a notebook with me in which I jot down keywords related to events, such as *arrest*, themes in conversations, such as *violence*, and my own emotions, such as *irritation*. At times, particularly in Durban, I used a recording device to record particular conversations and I always informed my informants when I used it. As it was practically impossible to record all 12 hours of each shift, I recorded conversations that I deemed to be important at the time. After the end of the day, or the following day, I wrote up field notes that I branded *mosaics of data*, which included jotted-down keywords, fragments of transcribed interviews, and detailed field notes written afterwards.

Unlike Susan Murray (2003), I do not distinguish between personal, observational or methodological notes, since I believe that such a categorization disregards the interconnectedness between emotions, method and knowledge. Rather, personal accounts and methodological issues are woven throughout the notes as a key ingredient of the empirical data, resulting in ‘messy texts’ (Denzin 1997) that voice various facets of the research process. In the coding of my notes, I only distinguish personal accounts when they are overtly explicit and at the forefront – otherwise, they blend in throughout the notes as an interwoven ingredient of the empirical data.

My means of writing notes is a very time consuming process and requires a lot of patience and perseverance. However, it pays off: when I read some of my

notes, I am able to mentally re-immersing myself back into that particular situation. Furthermore, writing up field notes is more than just recording data; it is the first step in data analysis. In my field notes, I often pose questions and try to identify patterns and connections between particular issues. In contrast to other disciplines and methods, data analysis is an ongoing process and does not only occur 'after' the data have been collected. In addition to my notes, I also maintain a detailed log of events, contact details and interview lists as a systematic way of organizing my data. This entire process greatly assists me when I analyse my data, as I have already started to use codes in these lists.

Another prominent technique that I employ throughout the research process is the frequent rereading of my field notes, both in and out of the field. This initially served as a means of recalling the research data, yet it has also allowed me to trace developments in my relationships with my informants and my analytical insights. Furthermore, as discussed by Kathleen DeWalt and Billie DeWalt (2002), the re-reading of notes is also a means in which researchers come up with new research questions. Unlike other methods, participant observation often demands adapting the research questions due to the nature of the collected data: our empirical data often steers us in a certain direction.

Longitudinal research

The fifth core element of participant observation concerns time: it is a method that needs to be carried out over a long period of time and requires a substantial amount of 'being there' (Becker 1970, in DeWalt and DeWalt 2002: 13). Previously, this entailed living in a field site (very often a village or neighbourhood) for a full annual cycle, in order to be able to document all of the events, changes and seasons of a year (Geertz 1988). For my PhD research, I spent 20 months in the field spread out across three years, of which one period was a bit longer than a year, and this is considered rather normal for many anthropologists.

Yet claims are increasingly made that long fieldwork periods are no longer necessary or possible. For example, due to personal circumstances, such as the difficulties of leaving one's family behind for longer periods of time, many anthropologists cannot stay in the field for months at end. Similarly, many research projects cover a period of three years, making it difficult to conduct a year of fieldwork. Therefore, in the comparative research project on security assemblages, I did not have the time to spend a year in each country, and my 12 months of fieldwork had to be spread across three field sites. This is also a reason why participant observation was not my main method in this project, and this is increasingly common in anthropology.

Yet at the same time, shorter periods of fieldwork are also criticised. John Van Maanen (2011), for example, discusses the emergence of '*Blitzkrieg fieldwork*', and states that 'formal interviews, sample surveys, focus groups, brief periods of observation ... replace lengthy in situ immersion' (Van Maanen 2011: 164). Such studies are critiqued for not acquiring that 'insider' perspective. The main reason behind this critique, and thus the perceived need for longer fieldwork periods, is

that more time means more immersion: research participants get more used to your presence and become more willing to discuss particular issues. Second, it takes time to acquire local knowledge, particularly the ‘tacit’ elements that often require a command of the local language, for example. The ability to speak the local language and/or dialect is identified by Kathleen DeWalt and Billie DeWalt (2002: 4) as a key element to participant observation. I concur, as not speaking Hebrew made it practically impossible to conduct participant observation in Israel. Whereas I accompanied security officers and attended training sessions in South Africa and Kenya, I was not able to do the same in Israel, and due to the nature and time frame of the project, learning Hebrew was not a possibility. This did not mean I could not conduct fieldwork, but that the opportunities for conducting participant observation were very limited.

Rapport

The issue of time is related to the sixth key element, namely establishing ‘rapport’, which is both a goal and a tool (DeWalt and DeWalt 2002: 40). Rapport is difficult to define, but it generally refers to a particular type of relationship between the researcher and the researched, whereby they share a common goal and feel comfortable with each other to communicate, so that the researcher can collect data from the researched. Trust is probably the most fundamental basis for rapport and establishing this takes time.

In my research, I have gained access to the companies by approaching the owners and/or managers and requesting permission to study their activities. Although my request to accompany security officers in their vehicles has sometimes been denied, primarily for (gendered) safety reasons, most owners have provided full consent for participation. Access is crucial to conducting such research and it is therefore extremely important that this is (partially) granted before one does the fieldwork. Yet rapport and building trust was primarily an issue with the armed response officers, as it took time for them to really open up to me. As has been the case for other researchers working on policing (Horn 1997; Marks 2004), many armed response officers suspected that I was a spy for company management, the police, or other companies, and that my intention was to ‘watch them’ and report back to a higher authority. Fortunately, this suspicion dwindled over time, and eventually the armed response officers came to trust me in both occupational and personal matters. In Kenya, where I had less time to conduct fieldwork, I attempted to gain credibility by providing copies of my PhD dissertation to interviewees. I hoped that, by showing I had researched the topic before, they would be more willing to discuss particular issues with me.

With some informants, rapport is established after a particular event or moment, regarded as a ‘break-through’, while with others it may be subtler and may develop slowly over time. With some informants, rapport entails a very friendly relationship, with whom I interact similarly to as I do with other friends outside the field. With other informants, rapport does not resemble a friendship, but is more professional and based on the research topic. For example, I regularly had lunch and

breakfast appointments with one company owner in Durban. Although we sporadically mentioned personal matters, such as family life, we primarily talked about the topic of my research (i.e. private security). It was not necessarily a very friendly and personal relationship, but it was one that was based on trust and comfort. Furthermore, I also established rapport with informants that I disliked, an issue that is not often discussed. Even though I disliked some of them, there was a certain amount of understanding between us and I was still able to collect fascinating data from them.

Spending long periods of time with people allows you to gain access to their perceptions and feelings. With armed response officers in Durban, I was able to collect data about their personal perspectives and experiences of doing security work. For example, discussing traumatic experiences is not common among armed reaction officers. When I enquired about particular traumatic episodes, common responses included: 'I can handle it', 'It's part of the job' and 'I've toughened up'. As time passed, however, several armed reaction officers opened up to me about their difficulties in dealing with the stress of the job. Through establishing rapport, I was able to know more about their personal lives, and many shared stories about the difficulties of 'turning off' after a day's work, not being able to communicate with their families, engaging in various forms of domestic violence, and having extramarital affairs to escape their 'failures at home'. Therefore, through developing rapport, researchers can incorporate the individual and personal experiences, which is a dimension that is often overlooked in policing studies. I argue that the emotions evoked while performing particular policing practices shape the nature of those practices and therefore the ontological nature of the individual performer. This reaffirms the need for ethnographic fieldwork that allows us to analyse emotions, perceptions, and multiple forms of meaning-making among our research participants.

Reflexivity

The seventh key element that is crucial to participant observation is reflexivity. The main critique addressed towards participant observation is that the researcher's engagement in the field eliminates objectivity. The idea is that the involvement of the researcher results in a 'clouded judgement', as the mere presence of the researcher alters the situation that the researcher is analysing. I concur that we as researchers influence our research setting. For example, elsewhere I discuss how I, as a woman, influenced the performance of masculine behaviour by armed response officers (Diphooorn 2013). Yet I am also convinced that my presence was not decisive, as many armed response officers exhibited this type of behaviour both before and after my fieldwork. Furthermore, I am certain that this does not prevent us from making a scientific analysis.

Yet in order to do this, researchers must engage in a process of reflexivity. Primarily due to the 'crisis of representation' and influence of feminist theory, the importance of reflexivity is currently undisputed. It is regarded as an 'unavoidable pre-condition' (Madden 2010: 23) to any ethnographic fieldwork

in order to ‘achieve a methodological rigour’ (Nilan 2002: 369). Yet although what is meant by *reflexivity* remains to be debated, there is a general consensus that ethnographers must understand their own position and role in the social world that they are studying and recognize that they, as persons with their own personality traits, backgrounds, and world perceptions, play a decisive part in shaping the entire research process (Coffey 1999; Denzin 1997; Madden 2010). I regard reflexivity as the ‘reciprocal interplay of one’s relationship with oneself and with others’ (Jackson 2010: 36). Furthermore, elsewhere I discuss the importance of reflecting on how our emotions influence how we participate in the field, i.e. ‘the emotionality of participation’ (Diphhoorn 2013). Yet this does not mean that we should all engage in ‘narrative ethnographies’ (Tedlock 1991) and produce ‘confessional tales’ (Van Maanen 2011) where researchers solely focus on themselves.

What people say and do

These seven elements constitute participant observation as a method and I want to highlight that participant observation is therefore more like a box of tools that conjunctively form a single method. It is therefore about the *combination* of the seven different elements. For example, participating in events without making field notes or observing a meeting without being reflexive is not participant observation.

Although there are numerous advantages, the main benefit of participant observation is that we not only analyse what people *say*, but we focus on what people *do*. The reality is that people often provide socially accepted answers during surveys and interviews. One of the best examples from my research in South Africa to exemplify this regards the use of violence. When I first started my fieldwork in Durban in 2007, I asked people during interviews about the use of violence, and specifically, whether armed response officers ever became involved in violent incidents. Almost everyone stated that violence was seldom used, that armed response officers were rarely instigators or victims of violent incidents, and they generally all frowned upon the use of coercion. I therefore initially assumed that violence would not be a central theme in my research. However, after a few months of fieldwork in 2008, it slowly became clear that violence is a large part of their everyday policing practices. Although violent encounters do not occur on a daily basis, they are certainly not rare, and as I argue elsewhere, violence is a central component of the occupational culture of armed response (Diphhoorn 2016). When I then discussed this with the same interviewees who had previously disregarded the centrality of violence, they proceeded to discuss at length about the problems associated with violence, recollected all sorts of stories where violence was used, and in some cases, even showed off their use of violence.⁴

A second example emerged during my fieldwork in Kenya where I focused on a formalized partnership between the Diplomatic Police Unit (DPU), a particular unit of the Kenyan Police, and several private security companies operating

in that area. This policing partnership comprises two main elements: joint-patrols whereby police officers patrol with the private security companies in their vehicles and the sharing of crime intelligence during monthly meetings. When I interviewed individuals involved in this partnership, they all stressed the formalized and organized nature of this partnership and described a structured system between the companies regarding the joint-patrols, with company A conducting patrols on Monday night, company B on Tuesdays, and so forth. However, when I eventually went on some of these joint-patrols, it became evident that much of this structure was non-existent. In fact, I realized that some companies were not conducting patrols at all, that some companies patrol on different nights and thus not on the same day every week, and that most of this was ad hoc and informal. Therefore, what was presented as a very formalized and structured policing partnership was in fact not the case. It was through participant observation that I discovered this discrepancy.

These differences are not simply glitches or disparities. Rather, I argue that they are fundamental insights into how non-state security operates. Knowing whether armed response officers engage in acts of violence and how they feel about it is crucial to understanding how security is related to crime, violence and fear. Knowing whether police officers and security officers patrol weekly within a structured system or on a very ad hoc basis, is important to know when analysing how state and non-state security actors interact with each other to shape the pluralised security landscape. It is therefore not simply about catching people in a lie, but it is about uncovering data that has theoretical implications. Through participant observation, I was able to discover particular practices and processes that determined my conceptual contribution to the scholarly work on private security. For example, my research in South Africa allowed me to develop the concept of ‘twilight policing’ (Diphooorn 2016), which refers to disciplinary, punitive and exclusionary practices that are performed in a twilight zone between state and non-state and are marked by uncertainty. Although I also used other methods, it was participant observation that generated the bulk of my empirical data and played the most fundamental role in my conceptual development. Furthermore, although my empirical data are based on South Africa and generalization is not my main aim, I argue that ‘twilight policing’ is a conceptual framework that can be applied to understand policing outside the South African context.

Concluding remarks

In this chapter, my main aim was to discuss the method of participant observation and how I used it in my research on private security officers and companies. In my discussion of what the key elements of participant observation are, I showed that participant observation is a very useful method to understand the daily practices and experiences of non-state security. Furthermore, participant observation concerns a combination of various methods that are interconnected, serves specific types of research questions and objectives, and demands a degree

of craftsmanship. Yet, like other methods, participant observation involves particular logistical and practical issues. For example, taking extensive field notes, learning the local language and spending long periods of time in a foreign place are physically demanding and time consuming activities. Furthermore, accompanying private security officers can place the researcher in dangerous situations, which inherently leads to a range of ethical and moral predicaments.

However, despite these disadvantages, I argue that participant observation provides tremendous insight into how non-state actors operate and is a very rewarding method. By spending long periods of time in a particular place, participating in a range of activities, observing what people do and the surroundings, continuously engaging in conversations and small talk, writing systematic field notes and being reflexive, we can access data about how people think and feel about security. Furthermore, as highlighted in the previous section, we are able to analyse not only what people say, but also what people do, and the potential incongruities between the two. Participant observation therefore provides data that other methods cannot, and when trained in how to use it and in combination with other methods, it acts as a truly significant method in researching non-state security providers.

Notes

- 1 For more information on this project, see www.security-assemblages.com.
- 2 Other scholars have made similar categorizations, such as Murray (2003), who differentiates between different roles in the field, and Hage (2010), who discusses the analytical, emotional, and political modes of participation.
- 3 This interview was conducted together with Francesco Colona, the PhD researcher of the comparative research project, who focuses on private security assemblages in Nairobi, Kenya.
- 4 This is also an example of how empirical data forces an alteration of the initial research question.

References

- Agar, Michael. 1996. *The Professional Stranger: An Informal Introduction to Ethnography*. 2nd edn. San Diego, CA: Academic Press.
- Buur, Lars. 2006. 'Reordering Society: Vigilantism and Expressions of Sovereignty in Port Elizabeth's Townships'. *Development and Change* 37(4): 735–757.
- Chisholm, Amanda. 2014. 'Marketing the Gurkha Security Package: Colonial Histories and Neoliberal Economies of Private Security'. *Security Dialogue* 45(4): 349–372.
- Coffey, Amanda. 1999. *The Ethnographic Self: Fieldwork and the Representation of Identity*. London: Sage.
- Denzin, Norman K. 1997. *Interpretive Ethnography: Ethnographic Practices for the 21st Century*. Thousand Oaks, CA: Sage Publications.
- DeWalt, Kathleen M., and Billie R. DeWalt. 2002. *Participant Observation: A Guide for Fieldworkers*. Walnut Creek, CA: AltaMira Press.
- Diphoorn, Tessa. 2013. 'The Emotionality of Participation: Various Modes of Participation in Ethnographic Fieldwork on Private Policing in Durban, South Africa'. *Journal of Contemporary Ethnography* 42(2): 201–225.

- Diphoorn, Tessa. 2016. *Twilight Policing. Private Security and Violence in Urban South Africa*. Berkeley: University of California Press.
- Driessen, Henk, and Willy Jansen. 2013. 'The Hard Work of Small Talk in Ethnographic Fieldwork'. *Journal of Anthropological Research* 69: 249–263.
- Evans-Pritchard, Edward Evan. 1940. *The Nuer: A Description of the Modes of Livelihood and Political Institutions of a Nilotic People*. Oxford: Clarendon Press.
- Feldman, Gregory. 2011. 'If Ethnography Is More Than Participant Observation, Then Relations Are More Than Connections: The Case for Nonlocal Ethnography in a World of Apparatuses'. *Anthropological Theory* 11(4): 375–395.
- Forsey, Martin Gerard. 2010. 'Ethnography as Participant Listening'. *Ethnography* 11(4): 558–572.
- Geertz, Clifford. 1988. *Works and Lives: The Anthropologist as Author*. Cambridge, UK: Polity Press.
- Ghassem-Fachandi, Parvis, ed. 2009. *Violence: Ethnographic Encounters*. Oxford: Berg.
- Goldstein, Daniel M. 2010. 'Toward a Critical Anthropology of Security'. *Current Anthropology* 51(4): 487–517.
- Goldstein, Daniel M. 2012. *Outlawed. Between Security and Rights in a Bolivian City*. Durham: Duke University Press.
- Gupta, Akhil, and James Ferguson, eds. 1997. *Anthropological Locations: Boundaries and Grounds of a Field Science*. Berkeley, CA: University of California Press.
- Hage, Ghassan. 2010. 'Hating Israel in the Field: On Ethnography and Political Emotions', pp. 129–154 in *Emotions in the Field. The Psychology and Anthropology of Fieldwork Experiences*, J. Davies and D. Spencer, eds. Stanford, CA: Stanford University Press.
- Higate, Paul. 2011. '"Cowboys and Professionals": The Politics of Identity Work in the Private and Military Security Company'. *Millennium – Journal of International Studies* 40(2): 321–341.
- Horn, Rebecca. 1997. 'Not "One of the Boys": Women Researching the Police'. *Journal of Gender Studies* 6(3): 297–308.
- Ingold, Tim. 2008. 'Anthropology Is *Not* Ethnography'. *Proceedings of the British Academy* 154: 69–92.
- Jackson, Anthony, ed. 1987. *Anthropology at Home*. London: Tavistock.
- Jackson, Michael. 2010. 'From Anxiety to Method in Anthropological Fieldwork: An Appraisal of George Devereux's Enduring Ideas', pp. 35–54 in *Emotions in the Field. The Psychology and Anthropology of Fieldwork Experiences*, J. Davies and D. Spencer, eds. Stanford, CA: Stanford University Press.
- Jensen, Steffen. 2008. *Gangs, Politics and Dignity in Cape Town*. Oxford: James Currey.
- Konopinski, Natalie. 2014. 'Borderline Temporalities and Security Anticipations. Borders'. *Etnofoor* 26(1): 59–80.
- Lee, Raymond M. 1995. *Dangerous Fieldwork*. Thousand Oaks, CA: Sage Publications.
- Lee-Treweek, Geraldine, and Stephanie Linkogle, eds. 2000. *Danger in the Field. Risk and Ethics in Social Research*. London and New York: Routledge.
- Liebling, Alison, and Betsy Stanko. 2001. 'Allegiance and Ambivalence. Some Dilemmas in Researching Disorder and Violence'. *British Journal of Criminology* 4 (3): 421–430.
- Madden, Raymond. 2010. *Being Ethnographic. A Guide to the Theory and Practice of Ethnography*. London: Sage Publications.
- Malinowski, Bronislaw. 1922. *Argonauts of the Western Pacific*. London: George Routledge and Sons.

- Manzo, John. 2004. 'The Folk Devil Happens to Be Our Best Customer: Security Officers' Orientations to "Youth" in Three Canadian Shopping Malls'. *International Journal of the Sociology of Law* 32(3): 243–261.
- Marcus, George E. 1995. 'Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography'. *Annual Review of Anthropology* 24: 95–117.
- Marks, Monique. 2004. 'Researching Police Transformation. The Ethnographic Imperative'. *British Journal of Criminology* 44: 866–888.
- Murray, Susan B. 2003. 'A Spy, a Shill, a Go-Between, or a Sociologist: Unveiling the "Observer" in Participant Observer'. *Qualitative Research* 3(3): 379–397.
- Nilan, Pamela. 2002. '"Dangerous Fieldwork" Re-examined: The Question of Researcher Subject Position'. *Qualitative Research* 2(3): 363–383.
- Nordstrom, Carolyn, and Antonius C.G.M. Robben, eds. 1995. *Fieldwork under Fire. Contemporary Studies of Violence and Survival*. Berkeley, CA: University of California Press.
- Pratten, David, and Atryee Sen, eds. 2007. *Global Vigilantes*. London: Hurst & Company.
- Punch, Maurice. 1986. *The Politics and Ethics of Fieldwork*. Beverly Hills, CA: Sage Publications.
- Rigakos, George S. 2002. *The New Parapolice: Risk Markets and Commodified Social Control*. Toronto: University of Toronto Press.
- Robben, Antonius C.G.M., and Carolyn Nordstrom. 1995. 'Introduction. The Anthropology and Ethnography of Violence and Sociopolitical Conflict', pp. 1–23 in *Fieldwork under Fire: Contemporary Studies of Violence and Survival*, C. Nordstrom and A.C.G.M. Robben, eds. Berkeley, CA: University of California Press.
- Rodgers, Dennis. 2006. 'Living in the Shadow of Death: Gangs, Violence and Social Order in Urban Nicaragua, 1996–2002'. *Journal of Latin American Studies*, 38(2): 267–292.
- Scheper-Hughes, Nancy, and Philippe Bourgois, eds. 2004. *Violence in War and Peace: An Anthology*. Oxford: Blackwell.
- Sluka, Jeffrey. 1990. 'Participant Observation in Violent Social Contexts'. *Human Organization*, 49(2): 114–126.
- Spradley, James P. 1980. *Participant Observation*. New York: Holt, Rinehart and Winston.
- Tedlock, Barbara. 1991. 'From Participant Observation to the Observation of Participation: The Emergence of Narrative Ethnography'. *Journal of Anthropological Research*, 47(1): 69–94.
- Van Maanen, John. 2011. *Tales of the Field: On Writing Ethnography*, 2nd edn. Chicago, IL: University of Chicago Press.
- Wakefield, Alison. 2003. *Selling Security. The Private Policing of Public Space*. Devon: Willan Publishing.

13 Using experimental methods in post-conflict countries to understand the effects of gender reforms in the Liberian National Police

Sabrina Karim

Introduction

The increase of experimental work in high profile journals as well as the launch of the new *Journal of Experimental Political Science* suggests a general trend of using experimental methods in political science, and of late, experimental methods in conflict/post-conflict countries.¹ The initial turn to experimental methods in political science began in the 1960s and 1970s when political scientists started to use behavioural theories to explain phenomena. In the 1980s, political scientists developed a social cognition framework for assessing political behaviour – political psychology – which also increased the use of experimental methods in the field. Around the same time, rational choice theory gained momentum and researchers turned to lab experiments to assess whether individuals behave in the rational ways theorized by game theoretic models (Druckman *et al.* 2006). The method grew in popularity, especially in the subfield of comparative politics (including American politics).

Much of the recent experimental work in political science has focused on assessing the impact of development programmes. The randomized controlled trials (RCT) – where the individuals being studied are randomly allocated one or more different treatments under the study – is considered the gold standard for programme evaluation in the medical field and in development economics, and is perhaps the most popular way to assess the effectiveness or the impact of particular policies, programmes and services. RCTs have also been used increasingly in post-conflict countries to assess peace building efforts and other programmes associated with security and development.

This chapter introduces the reader to a series of experimental work conducted with the Liberian National Police. My research has focused on two main questions. (1) How do civilian perceptions of and support for the government change when they are exposed to gender reforms in policing institutions in post-conflict countries? (2) How do group dynamics within the police institutions of post-conflict countries change due to gender reforms? The broader aim of the project was to understand how security sector reforms, starting with gender reforms in

the security sector, affect post-conflict countries. The literature on security sector reform in post-conflict countries is minimal; it is mostly relegated to Monica Toft (2010), which finds that the implementation of security sector reform leads to longer-term peace duration. While this book demonstrates the importance of security sector reform, it does not disaggregate reforms, and it does not provide a theoretical framework for evaluating reforms. Although my broader project assesses security sector reforms generally, experimental methods were also used to evaluate gender reforms in the security sector. The research mentioned in this chapter demonstrates why it was appropriate to use experimental methods, specifically RCTs, lab-in-the-field experiments and surveys to answer the questions and to begin to fill some of the gaps in the security sector reform literature. The studies discussed in the chapter were carried out between 2012 and 2015 in collaboration with the Liberian National Police.

The broader aim of this chapter is to describe when using experimental methods may be appropriate in studying non-state actors in post-conflict countries. Specifically, when are experimental methods appropriate instruments to use in conflict/post-conflict countries? Using experimental methods is not without challenges, such as ensuring compliance, safety, minimal spillover, feasibility and the need for vast resources. Yet, when such methods are used, they are often thought to be some of the most rigorous methods available to researchers. Thus, it is important to develop a set of criteria for when relying on experimental methods is appropriate. Referring to my experimental work with the Liberian National Police, the chapter provides theoretical and practical considerations for using experimental methods to research populations in conflict/post-conflict countries. It begins by introducing causal inference and providing an explanation of the use of various tools that employ causal inference. It then moves to explaining the advantages and drawbacks of using such tools. The chapter concludes with a discussion of the ethical implications of using experimental methods.

Addressing sources of bias in causal inference: experimental versus observational approaches

The use of experimental methods in political science comes from concerns about quantitative, observational studies' inability to address sources of bias. In general, there are three sources of bias from which quantitative observational studies – and to some extent survey work – may suffer: endogeneity, selection bias, and omitted variable bias.²

There are at least two potential non-experimental methods that researchers can use to answer the question: how do civilian perceptions of, and support for, the government change when they are exposed to gender reforms in policing institutions? In the first, using a quantitative observational method, the researcher collects data (ideally) from every post-conflict country about gender reforms in the security sector. For example, one may look at whether the post-conflict state instituted a quota for female officers or not.³ The researcher then uses statistical

analyses to assess whether post-conflict states that adopted a quota have higher rates of sexual violence in the country.⁴ The intuition is that with more women in the police force, individuals are more likely to report rape and domestic violence cases, and this is an indicator that there is increased levels of trust in the state because individuals are reporting cases to the state. In the second method, the researcher conducts a survey and asks whether or not individuals had contact with female police officers and then later on in the survey whether or not they trust the police force. Unfortunately, however, the three sources of bias affect both types of research designs.

First, endogeneity refers to the difficulty in assessing the causal direction of relationships. Statistical models are not able to account for loops in causality between independent variables (or the explanatory variables) and the dependent variable (the outcome of interest). In the observational, quantitative research design mentioned above, it is quite possible that states that experience high levels of sexual violence are more likely to adopt quotas for women in the police force. In the survey design, it is possible that those that trust female police officers seek out contacting female police officers. In both cases, it is not possible to tell which causal direction is at work.

Second, in technical terms, omitted variable bias refers to variables that are correlated with both the dependent and independent variables but that are absent from the analysis. Factors that may influence the relationships are not accounted for. For the quantitative, observational study, while it is possible to take into account observable factors such as the gross domestic product (GDP) of the country, regime type and other observable country-level variables, it is not possible to account for unobservable factors that might influence both whether states adopt female quotas and whether or not there are higher rates of sexual violence – i.e. the level of gender equality in the country. For the survey research design, it is possible to take into account observable variables such as sex, age, household income, etc., but unobservable variables, such as latent trust in the government, may still be missing from the analysis. Because researchers are unable to take into account every possible confounding factor, another method is required to get at cause and effect relationships.

Finally, selection bias/effect refers to the selection of individuals, groups, or data for analysis in such a way that proper randomization is not achieved. As a result, the sample obtained is not representative of the population intended to be analysed. In the quantitative, observational study, only post-conflict countries were selected. Post-conflict countries are not a random sample of countries in the world. This limits the ability to generalize whether or not quotas for women affect reporting rates of sexual violence globally. For the survey research design, it is possible that the researcher asked survey questions to only individuals that were at home – mostly females. The survey results may indicate that when the respondents, mostly female, had contact with female officers, they were more likely to report sexual violence. However, this conclusion is not valid, because men who should have had an equal chance of being selected were not included in the sample. It is possible that men may not have the same response as women

due to negative gender stereotypes about women or other patriarchal beliefs, but these potential results are not included in the findings.

The examples demonstrate the perils of using quantitative, observational studies and sometimes even survey designs to assess the effects of gender reforms in the security sector on citizens' perceptions and behaviour. The difference between observational studies (and surveys), which are common in the field of conflict studies, and experimental methods is the ability to randomize the causal factor or what may be called the treatment. The treatment is randomly assigned to a particular population defined by the researchers. Everyone in that particular population must have an equal chance of being selected to receive the treatment. Note that experimental studies focus on the random assignment of treatment to the subject – not to be confused with a random sample. For the former, each stage of the process in the research design ensures that all subjects have equal probability of receiving the treatment. A random procedure, whether a coin flip, a dice roll or a random number generator, determines whether the subject receives the treatment or not. Those that do not receive the treatment are considered a part of the control group. For the latter, subjects are chosen from a larger population randomly, whereby each subject is equally likely to be chosen. The process ensures that the observed and unobserved factors that influence outcomes are equally likely to be present in both the treatment and control groups. In other words, the groups are 'balanced'.

Random assignment is necessary for the proper assessment of causal relationships. However, to ensure that the results are generalizable to the broader population, a random sample from which to randomly assign treatment is also necessary. In other words, to ensure that the experiment has external validity, or that the findings can be applied broadly, random sampling is necessary to avoid problems of selection bias.

If we return to the original question about whether or not gender reforms increase reporting of sexual violence, then using experimental methods might ensure that the sources of bias mentioned above are avoided. I randomized police patrols by sex and then evaluated whether or not reporting of sexual violence increased (in addition to whether other outcome variables such as enhanced perceptions of the state/police forces or support for the government increased) in those areas that were patrolled by women compared with the areas that were patrolled by men, and compared with areas that were not patrolled at all (Karim 2016). Using this example, random assignment of treatment refers to the random selection of households to be visited by female police officers. Random selection refers to the random selection of villages selected for the entire study – these include the control villages where there will be no visits by policewomen and treatment villages. This approach ensures the preferred causal direction, ensures balance between the treated and untreated groups on unobservable characteristics, and ensures a random sample of the population of interest.

While it may seem as if experimental methods are quantitatively based, employing experimental methods may help bridge methodological divides between those who prefer case study and those who prefer large-*N* data analysis.

Due to the nature of field experiments, they can usually only be conducted in particular locations (one country, one region of a country, etc.). This means that researchers that use them must be intentional and employ rigorous methods for case design and case selection (Gerring 2008). Additionally, qualitative methods such as interviews (see Chapter 5 by Anja Mihr) and focus groups provide very helpful knowledge that lead to better designed experiments. They may provide important supplemental material that greatly contextualizes experimental results, particularly the mechanisms and processes through which treatments may have an effect (Paluck 2010).

Comparison of experimental methods

Field experiments

A field experiment refers to studies conducted in natural settings in which subjects are allocated randomly to treatment and control groups. Field experiments in post-conflict settings usually involve testing the impact of policies, in particular development programmes. They are also helpful in assessing the mechanisms through which certain outcomes occur. In the case of my research, for example, it was useful to test whether it was the sex of the officer or any visits, regardless of sex, by police officers that increased support for the state. I did this by sending male officers to randomly selected households and females to randomly selected households and no officers to another set of randomly selected households.

Other researchers have used field experiments to assess the impact of community-driven development programmes (Beath *et al.* 2012, 2013; Fearon *et al.* 2009, 2015) and land programmes (Blattman, Hartman, and Blair 2014). Moreover, field experiments in conflict and post-conflict countries served to test the impact of psycho-social programmes on ex-combatants (Blattman and Annan 2016) and of police patrols on trust in the government (Karim 2016; Blair, Karim, and Morse 2016).

There are several criteria that determine whether a researcher conducts a ‘true’ field experiment. According to Alan Gerber and Donald Green (2012: 10–11), these are whether the treatment used in the study resembles the intervention of interest in the ‘real’ world (authenticity), whether the participants resemble the actors who ordinarily encounter interventions (participants), whether the context within which subjects receive the treatment resembles the context of interest (context), and whether the outcome measures resemble the actual outcome of theoretical and practical interest (outcomes). When researchers implement field experiments, they should try to meet these criteria.

Authenticity

For authenticity, the intervention or the treatment should closely resemble what happens in the real world. For example, if the goal is to measure how police

patrols by women influence trust in the state, then the treatment should mimic how police patrols would normally be carried out by the state. Deviations from this approach may include having enumerators⁵ visit households and provide information about gender balancing reforms in the police. More generally speaking, ‘information treatments’ or surveys in which enumerators provide information about a particular policy and then assess whether that information changes peoples’ minds or behaviour about something, move the study away from being authentic. If experiments are not authentic, it may be hard to extrapolate what the findings of the experiment mean. In the latter scenario, it would be unclear if the actual reform or information about the reforms drive perceptual and behavioural changes.

Participants

In field experiments, it is important to get as close to the true population of interest as possible. Sometimes this may be difficult to achieve, especially if the research question is related to studying political elites such as high-level political figures or members of the security forces (Loewen *et al.* 2010). When research with the participants of true interest is not possible, then researchers have used a ‘convenience sample’ usually involving undergraduate participants or online systems such as Mechanical Turk (Berinsky *et al.* 2012). The downside to this is that it can be difficult to determine whether or not college students act similarly to the population of interest such as elites, police officers, or civilians in post-conflict settings.

Context

Another important consideration for field experiments is the environment in which the treatment is implemented. This includes whether the implementation of the treatment occurs in the location of interest. This may be difficult because it is often laborious to travel to post-conflict settings. Nevertheless, doing field work in the ‘real setting’ is important because it ensures that factors related to a different setting (such as a lab setting) do not influence civilians’ answers on surveys and behavioural outcomes. For example, if the study is on gender reforms in policing and the interactions with female police officers occur in a lab versus in the community, individuals may say they trust the female police officers when they actually do not.

Outcomes

The outcome measures should resemble the outcomes of interest in the real world. Researchers often use surveys to measure change in response to an intervention, but these only gauge changes in perceptions and not actual behaviour. While changes in perceptions are important to measure, the goal is (mostly) to understand whether the treatment induces some change in behaviour or a tangible

outcome of interest that improves the quality of life of the individuals. For that reason, looking at outcomes such as changes in reporting levels, income, community leadership positions, voting and other variables that actually capture changes in the behaviour or people's status better meets the criterion of outcome. Another concern is that sometimes researchers are interested in assessing concepts that are difficult to measure, such as 'legitimacy' and 'trust'. However, while surveys are often utilized to get at difficult concepts, researchers can sometimes employ creative methods to measure concepts such as legitimacy and trust.⁶

Potential disadvantages and trade-offs of field experiments

Although field experiments or randomized controlled trials may be useful for evaluating programmes, there are several downsides of using them. Field experiments require parsimonious theories (Gerber and Green 2012) because they usually allow researchers to test only a limited number of treatments or causal factors. The unit of analysis may be confined to the individual or at most to villages or certain organizations. Field experiments are not practical for theories of institutional change, explaining historical processes, and they are not practical for units of analysis that involve states, international systems, or fuzzy concepts such as power. Thus, the increased use of such methods could narrow the theoretical innovations in the field.

Additionally, outcome measures may be detectable in the short-term, but still not last in the long term. That is, the effects from programmes usually only last for a short period after the intervention and decay rapidly; this underlines the limited utility of field experiments for testing theories of (longer-term) institutional change on highly aggregated levels.

There are also concerns about external validity. Because experimental methods are resource-intensive and thus can only be implemented in limited settings (one country, one region of a country, etc.), the theory or causal mechanism(s) being tested may only be context specific. As a result, experimental methods may be helpful if researchers are assessing a novel causal claim to see if there is any merit to it. If there is support for the causal claim, then the onus of demonstrating the scope conditions for generalization is on the researcher. He or she should suggest other countries/areas where the theory may generate support and why it would or would not.

With limited resources to implement field experiments, researchers are forced to make difficult decisions that constrain the ability to include multiple parameters in the research design. In designing a field experiment, researchers must ensure that there are enough participants in the control groups and in all the different treatment groups. This is usually done by calculating the statistical power or the likelihood that a study will detect some effect when there is an effect to be detected. Highly powered studies ensure that the probability of concluding that there is no effect when, in fact, there is an effect (Type II error) decreases. Higher power is associated with a higher number of participants in the control

group and each treatment group or arm. Thus, the more treatment arms (i.e. causal factors) the study includes, the higher number of participants must be included in each group. Doing so raises costs, time and other resources.

Field experiments are expensive, time-consuming and high stakes. They often require that researchers receive large amounts of funding through grants. This limits the number of researchers that can engage in such efforts as there is a finite amount of available funding. It also means that field experiments are held hostage to budget cuts, funding cycles and grant guidelines. Field experiments usually take a long time to conduct, so they are not efficient for scholars who are pressed to publish frequently. They are also high stakes in that they are dependent on many outside factors that are beyond the control of the researcher. Often, researchers cannot closely monitor the implementation of the treatment and data collection process at all times because projects conducted in the field take a long time and researchers must rely on the assistance of locals or partner agencies to implement them. As a result, this sometimes leads to mistakes that jeopardize the integrity of the entire experiment, such as if proper randomization is not conducted. Field experiments are also susceptible to major shocks, which makes them high risk and costly if something goes wrong. As an illustration, researchers who were conducting field experiments in Sierra Leone or Liberia when Ebola broke out had to suspend their projects, which in some cases meant the project ended without being complete and funding was lost.

Lab-in-the-field experiments

Another type of field experiments are lab-in-the-field experiments. Lab experiments are often used by applied game theorists to test whether individuals behave rationally when making decisions and bargains. These games are often stylized portrayals of the real world, as participants are often ordinary people such as university students, and not legislators or diplomats or other populations of interest. A lab experiment becomes ‘in the field’ when it is conducted outside of a lab and in the actual location of interest. Lab-in-the-field experiments are carried out with actual representatives of the entity being studied. This is the method that I used to answer the second research question: how do group dynamics within the police institution change due to gender reforms?⁷ We (see Karim *et al.* 2016) selected over 600 Liberian National Police officers from the different police units and departments to participate. We then randomized the group composition by sex (no women, two women, four women, and six women) – the treatment was the number of women in each group and the control group constituted groups with no women. We then assessed the individuals’ and teams’ ability to do police work, as well as tested some standard behavioural games. Two typical games were used to assess whether the level of altruism or cooperation increases or decreases based on group composition.

- **Public Goods Game:**⁸ In the public goods game each player is allocated an endowment (i.e. generally corresponding to a consistent proportion of a

daily wage), and must decide how much of this endowment to personally retain and how much to put in a group pot. The total amount donated to the group pot is then doubled and redistributed evenly among all the players. According to game theoretic models that suggest that individuals are rational, for the individual, the optimal strategy is to give nothing and ‘free-ride’ on other players’ contributions to the collective fund, while the most profitable outcome for the group is that all players contribute everything. Variants of this game include telling people different scenarios for what will happen to the money or who will get the money.

- The Ultimatum Game:⁹ In the ultimatum game, the first player receives an endowment (i.e. generally corresponding to a consistent proportion of a daily wage) and proposes how to divide the sum between himself/herself and another player. The second player chooses to either accept or reject this proposal. If the second player accepts, the money is split according to the proposal. If the second player rejects, neither player receives any money. According to game theoretic models, the first player would give the lowest possible amount to the second person. Variants of this game involve saying who the second person is (someone in the same room or someone in different rooms).

The results of our study found that group composition does not affect groups’ level of cooperation or altruism, which runs contrary to some of the literature in social psychology that suggests that women are more cooperative and altruistic than men.

Survey experiments

Survey experiments include some sort of treatment that is written into the survey. Survey questions then evaluate how that particular treatment affected perceptions based on survey questions that come later (after the treatment). There are at least three types of survey experiments: ‘information’, ‘endorsement’ and ‘list’. The latter two are particularly helpful in post-conflict settings, where researchers may be interested in asking sensitive questions.

The use of informational triggers may influence how people make decisions. The treatment group receives some sort of information and the control group does not. Alternatively, the treatment group receives some variant of information while the control group receives some baseline information. The information may be provided in the survey (they read something or are told the information) or they may view a video that provides information. In my own research to assess whether actual visits by women or information about gender reforms are more important for enhanced perceptions and changes in behaviour, I conducted an informational survey experiment. The treatment included having the enumerator read information about gender reforms that have occurred in Liberia. Individuals in the control group did not get this information. The survey then assessed outcomes based on information or no information about gender reforms.

A second treatment may be to endorse a particular project or item on a survey and see if this leads to changed perceptions. Members of the treatment group are told that a particular policy is endorsed by a candidate or a group, while the control group is not told that the policy is endorsed by a particular candidate or group (Lyall *et al.* 2013). This allows the researcher to comparatively assess the level of individuals' support for non-state insurgent groups (such as insurgent groups), the state, or international actors.

In an example of an endorsement experiment, respondents were asked:¹⁰ '[THE LIBERIAN GOVERNMENT/THE TRADITIONAL LEADERS/SOME PEOPLE] say it is a crime to make children do work they are not able to do. How much do you personally agree?'

In this example, the researchers would assess the percentage of people based on whether they were randomly assigned the endorsement from the Liberian Government, traditional leaders or 'some people' (which is the control group). We thus learn how much individuals support traditional leaders over the state.

Finally, a list experiment includes statements about which participants may agree or disagree. The treatment group receives four statements while the control group receives three statements. The statements that the control group receives are usually statements that respondents are likely to all either agree or disagree on, whereas the one with four statements in the treatment group asks a question that is sensitive. The participants are not asked whether they agree or disagree with the statements, but rather just the number of statements they agree or disagree with. This allows the researcher to assess whether more people on average agree or disagree with the statement in question. Here is an example of such a list experiment:¹¹

I am going to present to you a list of four things [three, for the control group] that some people support and others do not. Please listen to these things and tell me HOW MANY you support. Do not tell me WHICH of these things you support, only how many of them you support.

The South American nations creating a central bank.

The assessment of a special tax to finance the expansion of the parks and green spaces in your neighbourhood.

The conservative ideology gaining more influence in the Colombian society.

The military forces having more freedom to defend the nation in the way they see fit [excluded for the control group].

There are many problems with implementing surveys as some questions of interest are sensitive; people are reluctant to answer questions that trigger difficult memories; that may cause trauma or that are embarrassing. Additionally, some groups of people may have some sort of bias in recalling events from the past. Such problems may lead to social desirability bias or non-random refusal. Social desirability bias is the tendency of survey respondents to answer questions in a manner that will be viewed favourably by others, particularly the enumerator or the group the enumerator represents. Non-random refusal is the tendency of survey respondents to choose to 'refuse to answer' in a systematic

way. Both are problematic because the researcher is not getting accurate responses for questions. These problems are exacerbated in conflict and post-conflict countries. Some questions on support for various insurgent groups or for the government may be sensitive and even dangerous to answer because insurgent groups are competing for support and if they hear that some individuals support one group versus another, it could jeopardize the responding individuals' lives. Nevertheless, both the endorsement and list experiments have been used in conflict and post-conflict countries to avoid these types of problems. (Bullock *et al.* 2011). There is evidence that both techniques produce substantively similar empirical findings even in places such as Afghanistan (Blair *et al.* 2014).

Ethical considerations in doing experimental research

This section considers the ethical implications of using the experimental methods mentioned above and suggests several considerations for researchers when using experimental methods.

At least in the US, all field experiments are subject to Institutional Review Board (IRB) authorization.¹² Because field experiments use human subjects, most researchers must gain approval from their home institution in the form of an IRB. But even if this is not mandatory, researchers should – as is the purpose of the IRB – ensure that all human subject research be conducted in accordance with federal, institutional and ethical guidelines. Each institution has a process for receiving IRB applications. Without IRB approval, researchers are often unable to get grant funding and are unable to publish their work. A key goal of IRBs is to protect human subjects from physical or psychological harm. Researchers are required to put together a research protocol that addresses the risks and benefits to subjects, the potential harm that can be done to subjects, the protection of subjects' identity, and demonstrate that participation is voluntary.

While the IRB is an important threshold for human subject safety, it is primarily designed to protect human subjects who participate in some form of medical study. In the social sciences, there are additional ethical considerations that researchers should consider before implementing experimental research designs. For example, the IRB does not consider the safety of enumerators nor of the researcher. Thus, especially when it comes to conducting field experiments, there are a number of other ethical guidelines to ensure that the researchers, enumerators and human subjects are unharmed from the experiment. Thus, the IRB is merely the minimal protection a researcher should take to ensure the work he or she does is ethical.

In addition to creating a protocol for the IRB, those involved in using experimental methods usually register their research design and create something called a pre-analysis plan. The pre-analysis plan usually includes the research design of the project, as well as the hypotheses that will be tested, and the empirical models that will be used to test the data. However, additional information could be included in the pre-analysis plan that includes decisions about enumerators, the researchers' control over the design and implementation of the

treatment, challenges and dangers associated with implementing the project, and provisions for information dissemination. These are all issues that relate to the ethics of implementing field experiments, but that are not addressed by the IRB.

When conducting field experiments, there may be a fine line with regards to how much control researchers have in manipulations that affect real people's lives. In a sense, researchers sometimes play the 'hand of God', choosing which people receive a treatment and which people do not and choosing what form the treatment will take. For example, in 2014, political scientists were reprimanded in Montana when they sent official-looking flyers to 100,000 Montana voters just weeks before Election Day. The manipulation could have swayed the election, and there was an official inquiry into the ethics of conducting such an experiment. While the 'do no harm' principle applies here, the question is more about whether researchers should be able to sway elections. How much involvement is too much involvement? Unfortunately, there are no generally applicable guidelines for such questions. Researchers must assess all potential outcomes – positive and negative – about their manipulation and then make an individual judgement call about whether it is worth conducting.

In my own research, I had to decide whether it was ethical for individuals who may have been traumatized by the war to receive visits by police officers who may have been perceived as threatening. The visits were justified because the government already had plans to do community policing in the rural counties of Liberia. My project enabled them to expedite this process. Thus, in this case, my project was actually able to help with state capacity building because it involved police officers becoming trained in community policing and it has potentially helped increase trust in the state.

One way to mitigate ethical problems is to conduct field experiments in partnership with local organizations. This means that the researcher is not the main decision maker when it comes to implementation of the treatment, but rather local organizations in consultation with the researcher make decisions about implementation and the local organization implements the treatment, not the researcher. This ensures that local context and local knowledge about the potential negative and positive impacts are taken into consideration. It also better guarantees that the goal of authenticity and context are achieved. Working with local partners helps to ensure that a treatment is something that the group or the government would actually do. Even though the researchers have less control over the implementation, it is important to implement something that would have been implemented by a campaign, by the government, or NGO even without the involvement of the researcher. The researcher 'only' ensures that the process of receiving the treatment is random.

Still, one criticism of researchers' control over randomization has been that the process by which individuals receive treatments is not organic. Rather, researchers choose who gets the treatment and who does not. This means that if there is an advantage to the treatment (i.e. it increases household income), then only select members of the population will get the benefit. However, researchers have argued that the process for selection is often more fair than the ones that

exist or that occur organically. Because the process is random, everyone has an equal chance of receiving the treatment, which may not be the case otherwise.

There are also potential challenges and dangers of doing fieldwork for the researcher. Conducting field experiments in conflict and post-conflict settings makes researchers susceptible to particular dangers such as diseases, breakdowns in law and order, crime, road accidents, and even sexual harassment. In my case, I had to delay my project for one year due to the Ebola epidemic in 2014, but the delay strengthened my project and I developed a much more rigorous memorandum of understanding with the research team, which included what to do if there was another outbreak. Researchers may do well to think about the potential dangers and challenges and ensure that they have the resources, insurance and contingency plans for conducting field work in such environments. One recommendation is to do an initial visit to the location before actually conducting the experiment in order to understand the challenges associated with working in the country. This also ensures that the researcher has a greater understanding of the context of the setting for the field experiment, which may better inform how the experiment will be carried out. It also contributes to (re-)considering the feasibility of the research design.

Conducting field experiments requires the use of local enumeration teams. This means that, in addition to considerations about the researcher's safety, the researcher must think about the safety of the enumeration team as well. Regardless of how the data collection team is organized, the researcher and the enumeration team should create a memorandum of understanding about the conditions under which the project will be implemented. The parameters for this should include what happens if researchers get sick on the job, what types of crisis may put an end to the project, which communities are safe and accessible enough for project implementation, what happens if, upon entering a particular area, it is actually unsafe and enumerators are unable to go there, and how much communication the enumeration team will have with the researcher.

For the data collection, while many countries have large survey firms that are often utilized to implement surveys and field experiments, there are several benefits and challenges to using such firms versus creating and training a team. Large survey firms will have already done feasibility assessments and will be able to advise researchers about their limitations. They are also likely to have the personnel, equipment and infrastructure to carry out the project. Thus, for those that have minimal experience in a country, such firms may be the best option. However, they tend to be very expensive and the researcher may have very little control over the process. This could lead to poor quality of data collection, as it becomes difficult for the researcher to ensure that the proper protocol and procedure was utilized.

Another option may be for the researcher to invest time and energy in creating his or her own enumeration team.¹³ This may be possible through using local college, Masters or PhD students. These students may be eager to learn about research methodology and, in employing them, the researcher may actually be contributing to skills development in the country, which is particularly useful for individuals in post-conflict countries. The researcher could spend time conducting training and building a relationship with the enumerators. The enumerators may

then be able to provide valuable information about the conditions under which the project may be implemented and other information that goes into the memorandum of understanding. The drawback of this approach is that it is more time consuming. However, the benefit is that the researcher may have more control over the implementation of the process, which means there may be less ‘shirking’.

A final, important ethical consideration is about how much the communities benefit from the research and how much they know about the research and the results. Perhaps one of the most important parts of conducting field experiments is conveying the information back to the communities in which the experiments occurred. This ensures that the research process is two sided and that both the researcher and the communities gain in knowledge (Kaplan 2015). Often times, researchers do not have enough resources allocated in their budgets to be able to give back in this way. Thus, it may be the duty of organizations that provide grants to incorporate a section in grant proposals requiring researchers to explain how they would disseminate their research to the participants and an appropriate budget for doing so. In addition to, or instead of, information dissemination, researchers may find other ways to give back to the community or their partners. For example, they may provide some sort of training for their enumerators, for their partner organization, or even in communities. Regardless, researchers that implement field experiments may do well to think more about the imprint they leave, and how the knowledge that they have may be made intelligible to the local population. For example, when I conducted a survey in 2012, I presented the results to the community during a community forum by having enumerators act out the results.

Conclusion

The chapter has provided an account of the growing use of experimental methods in conflict and post-conflict countries. The growing trend has necessitated a set of criteria by which researchers can evaluate whether or not to use experimental methods in post-conflict countries. Based on the above discussion, the following is a summary of questions researchers should ask themselves before using such methods with non-state actors in post-conflict countries. If the answer is yes to the questions, then experimental methods may be appropriate.

- 1 Is my research subject to high levels of bias if conducted through analysing observational data/surveys (endogeneity, omitted variables and selection bias)?
- 2 Is my research question limited to a unit of analysis that is the individual or groups?
- 3 Is my research question limited to only one or a few mechanisms?
- 4 Is this a new theory or causal mechanism being tested?
- 5 Are my scope conditions well-defined for the theory/causal mechanism?

The closer the research design is to authenticity, ‘real’ participants, ‘real context’ and ‘real’ outcomes, the more closely it aligns with a ‘true’ field experiment. However, there may be certain instances where a lab-in-the-field experiment

may be more appropriate, such as if researchers are interested in testing whether individuals behave according to outcomes predicted from game theoretic models – do they behave rationally – given some sort of treatment (such as varying group composition by sex). In other cases, survey experiments may be more appropriate due to the danger of conducting actual field experiments in post-conflict countries, or because individuals may not answer some questions truthfully. Thus, if researchers are trying to elicit answers to questions that are particularly sensitive, survey experiments may be a useful method.

Finally, other criteria relate to more practical and ethical aspects of using experimental methods. Practically, researchers need resources and time to use experimental methods. Additionally, there are certain ethical challenges that must be considered when using these methods. Researchers should work with partner organizations, choose enumeration teams with care, ensure that there is a memorandum of understanding with all participants in the study, and describe how researchers intend to convey information back to communities or think of other ways to give back to researched communities. Regardless of these suggestions, much more work should be devoted to setting guidelines for social scientists that wish to employ these methods and especially for those that wish to employ them in conflict and post-conflict environments, where the risks and stakes are higher.

Despite the challenges, field experiments, lab-in-the-field experiments and survey experiments represent new tools in a big tool box of methods that researchers studying conflict and post-conflict countries can use, and the use of such methods has the potential to create new knowledge that will be valuable to the field.

Notes

- 1 For a background of experimental work in International Relations, see Hyde (2010).
- 2 For in-depth explanations of these biases, see King *et al.* (1994).
- 3 In Liberia, the government, with help from the UN Mission in Liberia (UNMIL), instituted a 15 per cent quota for women in the police force in 2005, a 20 per cent one in 2008, and a 30 per cent one in 2012.
- 4 Other outcome variables may be measured, but here higher levels of sexual violence indicates higher reporting rates, which is an indicator of increased support for the state.
- 5 Enumerators are research assistants who are from the country in which the experiment is occurring, and who help carry out the experiment or surveys.
- 6 For example, see Blair (2015) and Karim (2016). In the latter study, to measure support for the government (such as inclination to start paying taxes), researchers may give individuals compensation for taking a survey and then asking them if they want to donate some proportion of what they received to the state to improve security services.
- 7 See Karim *et al.* (2016), Blair, Karim, Gilligan, and Beardsley (2016). For applications in post-conflict countries related to peacekeeping and trust and legitimacy, see Blair (2015).
- 8 You are with a group of people when a woman approaches with a proposition. She offers each person \$5 in one-dollar bills and says that you can each keep some or give to the pot. She says if you give to the pot, that amount will be doubled and then split among everyone. What do you do?

- 9 You are standing on the sidewalk with a friend, minding your own business, when a man approaches with a proposition. He offers you \$10 in one-dollar bills and says you can keep the money, under one condition: You have to share some of it with your friend. You can offer your friend as much or as little as you like, but if your friend rejects your offer, neither of you get to keep any of the money. What do you do?
- 10 Taken from Blair, Karim, and Morse (2016).
- 11 Taken from 'Controlling Civilians? Examining Support for the Military in Colombia' (Matanock and Garcia-Sanchez 2015). The article compares using direct questions versus list questions to gauge truthful answers.
- 12 The IRB is a committee, usually housed within an academic institution, which has been formally designated to approve, monitor and review biomedical and behavioural research involving humans.
- 13 This is the approach that I took. I have worked with the same enumeration team for every project and have made it a point to do some kind of training (i.e. on randomization, data analysis, grant writing, etc.) each time I go to Liberia.

References

- Beath, Andrew, Fotini Christia, and Ruben Enikolopov. 2012. Winning Hearts and Minds through Development? Evidence from a Field Experiment in Afghanistan. The World Bank. Policy Research Working Paper Series. <http://ideas.repec.org/p/wbk/wbrwps/6129.html> (accessed 8 February 2014).
- Beath, Andrew, Fotini Christia, and Ruben Enikolopov. 2013. 'Empowering Women through Development Aid: Evidence from a Field Experiment in Afghanistan'. *American Political Science Review* 107(03): 540–557.
- Berinsky, Adam. J., Huber, Gregory. A., and Lenz, Gabriel. S. 2012. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3): 351–368.
- Blair, Robert. 2015. *Four Essays on Peacebuilding and the Consolidation of State Authority*. PhD Dissertation. Yale University.
- Blair, Graeme, C. Christine Fair, Neil Malhotra, and Jacob N. Shapiro. 2013. 'Poverty and Support for Militant Politics: Evidence from Pakistan'. *American Journal of Political Science* 57(1): 30–48.
- Blair, Graeme, Kosuke Imai, and Jason Lyall. 2014. 'Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan'. *American Journal of Political Science* 58(4): 1043–1063.
- Blair, Robert, Sabrina Karim, Benjamin Morse. 2016. 'Building Trust in a Reformed Security Sector: A Field Experiment in Liberia'. International Growth Centre, June. Available from: www.theigc.org/wp-content/uploads/2016/08/Blair-et-al-2016-Policy-Brief-v3.pdf (accessed 21 October 2016).
- Blair, Robert, Sabrina Karim, Michael Gilligan, and Kyle Beardsley. 2016. Policing Ethnicity: Lab-in-the-Field Evidence on Discrimination, Cooperation and Ethnic Balancing in the Liberian National Police. Working Paper.
- Blattman, Christopher, and Jeannie Annan. 2016. 'Can Employment Reduce Lawlessness and Rebellion? A Field Experiment with High-Risk Youth in a Fragile State'. *American Political Science Review*, 110(1): 1–17.
- Blattman, Christopher, Alexandra C. Hartman, and Robert A. Blair. 2014. 'How to Promote Order and Property Rights under Weak Rule of Law? An Experiment in Changing Dispute Resolution Behavior through Community Education'. *American Political Science Review* 108(1): 100–120.

- Blattman, Christopher, Julian Jamison, Eric Green, and Jeannie Annan. 2014. 'The Returns to Cash and Microenterprise Support Among the Ultra-Poor: A Field Experiment'. SSRN Electronic Journal. Available from: www.ssrn.com/abstract=2439488 (accessed 10 October 2015).
- Bullock, Will, Kosuke Imai, and Jacob N. Shapiro. 2011. 'Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan'. *Political Analysis* 19(4): 363–384.
- Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia. 2006. 'The Growth and Development of Experimental Research in Political Science'. *American Political Science Review* 100(4): 627–635.
- Fearon, James D., Macartan Humphreys, and Jeremy M. Weinstein. 2009. 'Can Development Aid Contribute to Social Cohesion after Civil War? Evidence from a Field Experiment in Post-Conflict Liberia'. *American Economic Review: Papers & Proceedings* 99(2): 287–291.
- Fearon, James D., Macartan Humphreys, and Jeremy M. Weinstein. 2015. 'How Does Development Assistance Affect Collective Action Capacity? Results from a Field Experiment in Post-Conflict Liberia'. *American Political Science Review* 109(3): 450–469.
- Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton & Company.
- Gerring, John. 2008. 'Case Selection for Case Study Analysis: Qualitative and Quantitative Techniques', pp. 645–684 in *The Oxford Handbook of Political Methodology*, J.M. Box-Steffensmeier, H.E. Brady and D. Collier, eds. Oxford, UK: Oxford University Press.
- Hyde, S.D. 2010. 'The Future of Field Experiments in International Relations'. *The ANNALS of the American Academy of Political and Social Science* 628(1): 72–84.
- Loewen, Peter. J., Daniel Rubenson, and Leonard Wantchekon. 2010. Help Me Help You: Conducting Field Experiments with Political Elites. *The ANNALS of the American Academy of Political and Social Science*, 628(1), 165–175.
- Lyall, Jason, Graeme Blair, and Kosuke Imai. 2013. 'Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan'. *American Political Science Review* 107(4): 679–705.
- Kaplan, Oliver. 2015. "Taking it to the Streets: Engaged Research in Political Science," ISA Annual Conference, New Orleans.
- Karim, Sabrina. 2016. *Finding the Right Security Sector Strategy: The Goldilocks Problem in Post-Conflict States*, PhD Dissertation, Emory University.
- Karim, Sabrina, Kyle Beardsley, Michael Gilligan, and Robert Blair. 2016. 'International Interventions to Empower Women in Post-Conflict Countries: Lab-in-the-Field Evidence from the Liberian National Police'. Working Paper.
- King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Matanock, Aila M. and Garcia-Sanchez, Miguel. 2014. Does Counterinsurgent Success Match Social Support? Evidence from a Survey Experiment in Colombia. APSA 2014 Annual Meeting Paper. Available at SSRN: <https://ssrn.com/abstract=2453046>
- Paluck, Elizabeth Levy. 2010. 'The Promising Integration of Qualitative Methods on Field Experiments'. *The Annals of the American Academy of Political and Social Science* 628: 59–71.
- Toft, Monica D. 2010. *Securing the Peace: The Durable Settlement of Civil Wars*. Princeton: Princeton University Press.

14 Empirical assessment of (policy) effectiveness

The role of business in zones of conflict

Melanie Coni-Zimmer and Klaus Dieter Wolf

Introduction¹

Transnational corporations operate in conflict zones around the world. Most prominent are cases where they have become entangled in conflict dynamics or have supported authoritarian regimes. Often cited examples have been companies that buy ‘conflict diamonds’ from African countries or, more recently, that export ‘conflict minerals’ from the Democratic Republic of the Congo (DRC). In contrast, corporate social responsibility (CSR), a concept that has received increased attention since the 1990s (Coni-Zimmer 2014), emphasizes that companies can and should go beyond simply doing business by engaging in CSR and governance.² This discourse has spurred research on the potentially positive role of business and its contributions to governance in zones of conflict (Wenger and Möckli 2003; Haufler 2010; Deitelhoff and Wolf 2010b).

This chapter addresses the challenges of conceptualizing and measuring the effectiveness of corporate engagement in zones of conflict. It mainly draws on the experience the authors gained during a collaborative research project on the role of business in zones of (violent) conflict conducted at the Peace Research Institute Frankfurt.³ The research question examined in this project was: under what conditions do corporations contribute to governance relevant to peace and security in zones of conflict (Deitelhoff and Wolf 2010a)?⁴

The project was developed in response to a biased state of the research that has emphasized the negative role of business; that is, how corporations contribute to conflict rather than to peace and security. The increasingly important role of non-state actors has become a major issue in global governance research in the fields of political science and International Relations. In general, the call for an increased engagement of non-state actors rests on the assumption that they make governance more effective and legitimate (Reinicke and Deng 2000; Dany 2013: 81). Although research has highlighted the positive role of civil society organizations and their contribution to governance, the role of business actors has been neglected for some time. In addition, the role of non-state actors has been broadly discussed in areas considered ‘low politics’, such as environmental politics and development, but it has hardly been considered in the field of security, which is often perceived as the prerogative of the state. As a result, our

knowledge about the role of business actors in the areas of peace and security is limited. But given the growing discourse about CSR, it was reasonable to assume that businesses also have the potential to contribute to governance in zones of conflict. Moreover, there was a gap in research with regard to explaining different forms and degrees of corporate engagement in conflict zones. Consequently, our aim was twofold: to generate more systematic, descriptive evidence of corporate engagement in conflict zones, and to provide explanatory knowledge about the conditions under which corporations show varying degrees of engagement. If the engagement of business actors in conflict zones could indeed be shown to be effective, this would constitute an important argument for engaging this group of non-state actors in governance processes.

The distinction of corporate governance contributions, that is, the dependent variable in the project, was based on the concept of effectiveness as developed in regime theory (see below). Governance contributions were conceptualized in terms of *output* and *outcome* effectiveness (see Wolf *et al.* 2007; Deitelhoff and Wolf 2010a: 11). It was furthermore defined that these contributions had to be both intentional and voluntary – that is, they were not merely a by-product of business activities and they had to go beyond what companies were required to do by law.⁵ A distinction was made between governance-related activities that directly address the level of violence in conflict zones (security governance) and contributions that indirectly address the underlying causes of violent conflict, such as human rights and corruption, the socio-economic dimension of conflict or socio-cultural problems (Feil *et al.* 2008: 6–7). Direct contributions to security governance might include corporate standards for dealing with public and private security forces or contributing to security sector reform. For example, oil companies such as Shell implement the Voluntary Principles on Security and Human Rights that regulate their dealings with public and private security forces with the aim of preventing human rights violations and enhancing the human security of communities (Zimmer 2010: 65). Indirect contributions might include programmes to support civil society, the introduction of corporate human rights standards or the implementation of community development programmes. The latter nowadays constitute a standard practice of all major extractive corporations with large operations on the ground (Jacobs 2015; Zimmer 2010: 69–71).

The conceptualization of governance contributions based on the distinction between output and outcome effectiveness, as well as its pros and cons, will be discussed in the second section of this chapter. We argue that focusing on the two dimensions of output and outcome effectiveness is a sensible middle ground that seeks to reconcile two criteria: political relevance and the accessibility of data. In the third section, we will briefly describe the challenge of case selection and discuss the research methods used to collect the data. Method triangulation based on document analysis and interviews with companies and stakeholders (see also Chapter 6 by Joakim Berndtsson), as well as field research, allowed the researchers to gain valuable insights into the output and outcome dimensions of corporate governance contributions. The chapter concludes with a discussion of some of the broader implications of using this conception of effectiveness.

The effectiveness of corporate engagement⁶

Effectiveness can be conceptualized in many different ways,⁷ and analysing effectiveness has been an important strand in research on regimes. Regimes are usually defined as sets of ‘principles, norms, rules and decision-making procedures around which actors’ expectations converge’ (Krasner 1983: 2). Relevant research has focused primarily on intergovernmental regimes, such as the nuclear non-proliferation regime or various environmental regimes. When analysing the effectiveness of regimes, researchers have often distinguished among outputs, outcomes and impacts (see Easton 1965; Young 2004).⁸ ‘Outputs’ refer to policies, regulations and the development of instruments put in place as a response to a regime; ‘outcomes’ relate to measurable behavioural changes of regime members; and ‘impacts’ are ‘contributions regimes make to solving the problems that led to their creation in the first place’ (Young 2004: 12–13). It is fair to say that most research has focused on the output and outcome dimensions of effectiveness and not on the impacts and other broader consequences of regimes, owing mainly to difficulties in measurement and in attributing causality (Underdal 2004). More recently, the distinction among these three dimensions has also been used to analyse the effectiveness of public–private regimes (see Huckel *et al.* 2007; Flohr *et al.* 2010; Beisheim and Liese 2014). In our research project, this approach was adapted⁹ and used to analyse corporate engagement specifically in conflict zones (see Table 14.1).

Corporate output

Corporate output refers to the (self-)commitments of actors or, simply put, corporate policies. Of the three dimensions of effectiveness, output is the most manageable in scholarly terms, but it is also the least interesting in practical terms. Business actors’ self-commitments are often suspected of being only window dressing, unrelated to their real intentions or behaviour, or of not contributing to solving pressing societal problems (Vogel 2005).

However, analysing the output dimension sheds light on whether corporations have clearly articulated policies towards security in a conflict zone and whether

Table 14.1 Dimensions of effectiveness

<i>Dimension of effectiveness</i>	<i>Focus of attention</i>	<i>Analytical accessibility</i>	<i>Political relevance</i>
Output	Corporate (self-) commitments	Very high	Low
Outcome	Corporate behaviour	Moderately high	Substantial
Impact	Corporate contribution to problem-solving	Low	High

Source: own table.

they have goals when it comes to more structural causes of violence. Therefore, one can more explicitly investigate the intentionality of security-related corporate engagement and goals by examining output rather than the other dimensions. The output dimension is also where the 'design evaluation' can take place (e.g. by addressing the question of how strong the monitoring mechanisms are). Here, one can study the overall scope of commitments (global, national or local), as well as different actor constellations (unilateral, multilateral corporate commitment or multi-actor/sectoral commitment).

Output is relatively easy to address in analytical terms. One might look at the statements made by individual companies or at collective business or multi-stakeholder commitments that can be found, for example, in codes of conduct or sustainability reports (see below). Empirical results for different outputs can be measured and evaluated along an axis that ranges from proactive self-commitments to no self-commitments at all. To give an example, a company operating in conflict zones might have an individual statement referring to the use of and relationship with public and/or private security forces, or it might be part of a global standard initiative such as the Voluntary Principles on Security and Human Rights (see Freeman and Hernández Uriz 2003). The company thereby acknowledges a certain responsibility for protecting and promoting human rights and commits itself to following certain principles and implementing certain measures to prevent human rights violations by security forces.¹⁰

Corporate outcomes

Focusing the evaluation of effects on analysing outcomes will provide a good balance between being interesting (i.e. politically relevant) and manageable (in analytical terms) at the same time. It goes beyond analysing formal commitments in statements or codes of conduct (which are just 'words'); instead, it addresses and evaluates companies' behaviour and their compliance with their own commitments in zones of conflict: do corporate self-commitments indeed change corporate behaviour? Do companies actually comply with the (normative) standards to which they have committed themselves?

Analysing the outcome dimension of corporate engagement is more ambitious than assessing output effectiveness; however, when compared with impact assessments (see below), it is often still more practicable. At any rate, it requires looking beyond corporate statements of policies. Collecting data for analysing corporate behaviour should not rely on corporate self-reporting but should include other sources, such as interviews, observation and/or reports from stakeholders.

Different outcomes can be measured and evaluated along the axis from no corporate activities related to peace and security (limited to just doing business) to proactive activities that compensate for the failure of host states to provide security and related public goods. To give an example, from our research project, a company that has committed itself to the Voluntary Principles on Security and Human Rights

takes active steps to implement the standard by conducting risk assessments and offering human rights training for security personnel in conflict zones.

Impact

Impact (or problem-solving effectiveness) relates to the question of whether policies and activities that have been developed have contributed to mitigating or resolving real-world problems, such as the violent conflicts or human rights violations that led to their development in the first place. Of the three effectiveness dimensions, impact is clearly the most demanding in an analytical sense, but it is also the most interesting in terms of political relevance. Impact is ‘the ultimate concern of decision-makers and stakeholders’ (Underdal 2004: 34), and attributable impact would doubtless offer the strongest argument for bringing the private sector into the realm of conflict prevention and peacebuilding.

However, manifold analytical problems arise when one attempts to apply the impact yardstick of success. They start with a simple question: impact on what? It would be most interesting to know whether positive corporate engagement had an effect on the level of physical violence. However, security-related impacts might also include conflict-specific ‘correlates of peace’, which may affect the level of physical violence indirectly, as in creating a more stable and better governed political environment (e.g. promoting the rule of law, ensuring human rights and reducing corruption) or a more sustainable social, economic and ecological environment (Feil *et al.* 2008: 6–7). For example, implementing the Voluntary Principles on Security and Human Rights could be conceptualized as having an impact either on the human rights situation of the local population or on the level of violence.

Even greater challenges to measuring the impact of corporate contributions result from the problems of multi-causality and counterfactuals (for an overview, see Underdal 2004; see also Section II in this volume). How can we control for other factors that might explain changes in the level of violence in conflict zones?¹¹ What difference would the absence of a given private engagement make, and how can we discern it? Even if we observe that a corporation implements the Voluntary Principles on Security and Human Rights (and/or other measures) and that the level of violence has decreased in a given area, these two phenomena may not be causally linked; the reduction in the level of violence might be due to specific measures taken by state or other actors or to a general improvement in the relationship between the parties to the conflict. Moreover, in many conflict regions, there are usually simultaneous interventions by a variety of actors – state and non-state – making it difficult, if not impossible, to disentangle the impact of a certain intervention.

Given these analytical problems, and having to choose between identifying and assessing corporate engagement (output and outcome) or its impact, the former is clearly less difficult to translate into operational tools.

Based on these considerations, we chose to focus our research on the output and outcome dimensions of corporate engagement. Caught in the dilemma

between political relevance and analytical accessibility, measuring outcome may be less relevant politically than the impact of corporate activities in the sense of its proximity to an 'ultimate' success. In favour of focusing on outcome, it can be argued that outcome is closer to impact than mere output is and may even be a precondition for impact. Outcome also is less difficult to observe and measure than is impact. Unlike impact, outcome offers a yardstick for measuring success that avoids the complex methodological problems of counterfactuals and multi-causality, which are inseparably linked to studying and evaluating impact. Admittedly, the framework suggested is a pragmatic one, in that it values analytical accessibility more than immediate political relevance. Its pragmatism may have the advantage of being equally appealing to both scholarly research and practical, policy-oriented analysis. The choice to focus on output and outcome has a number of implications for data generation that will be discussed in the next section.

Case selection and data: combining document analysis, interviews and field research

Before we address the process and challenges of generating data to measure corporate output and outcome in more detail, we will first discuss case selection, because it poses specific problems when the focus is on business actors.

Case selection

There is a wealth of political science and International Relations literature on how to select cases (e.g. George and Bennett 2004; Gerring 2008). Generally speaking, in qualitative research, case selection is not random but is guided by the researchers' theoretical interests (Gerring 2008: 645–646). In our project, however, it was complicated by the collaborative nature of the research project, in which the studies conducted by all the participating researchers were intended to contribute to the project's overall aims; that is, describing and explaining corporate engagement in conflict zones.¹² This approach required a relatively high level of coordination in order to find a compromise between top-down considerations (related to the overarching goals of the project) and bottom-up considerations (related to the goals of the individual researchers).

For the overall research project, company cases were chosen that reflected different company and production characteristics; different political, social and market environments; and different conflict characteristics. These sets of factors were all deemed to be potentially relevant in explaining corporate behaviour (see Wolf *et al.* 2007; Deitelhoff and Wolf 2010a: 19). Project members jointly developed a codebook that included guidance on how to analyse the dependent and potential explanatory variables, including working conjectures that were sometimes contradictory. Among other studies, the research project included a study on the oil industry in Nigeria (Zimmer 2010), one on the tourism industry in Israel and Palestine (Fischer 2011) and one on beverage companies in

Rwanda and the DRC (Feil 2012). Case selection involved industry sectors that were usually covered in the literature and for the most part were perceived as fuelling conflict (e.g. the oil industry), as well as others that had been hitherto neglected (e.g. beverage companies).

The first criterion used in selecting cases was the existence (or only recent termination) of a violent conflict in a certain country or region.¹³ We then identified cases with a high likelihood that varying forms of corporate engagement could be observed. For example, in a study on the role of the oil industry, the first step was to identify countries where both violent conflicts and an operating oil industry were present. The second step was to compile lists of oil and gas companies operating in these countries. In this context, one particularity when doing research on non-state actors in conflict zones is the often potentially large number of cases. The information necessary to compile such lists was sometimes easily accessible via corporate websites, industry associations or NGO reports; but, more often, it was difficult to obtain, as when smaller or local companies rather than larger, transnational companies were investigated.

The final selection of the country and company cases took into account data richness, the potential availability of companies with different characteristics and the accessibility of the conflict zone. For example, Nigeria was selected as a conflict zone for investigation, because many important oil companies are active there. Shell, Statoil and ExxonMobil were chosen as the company cases, because this selection of cases allowed us to analyse the influence on corporate engagement of the different political and societal environments in the corporations' home states (Zimmer 2010). At the same time, cases were selected in close coordination with other researchers involved in the project. Examining the influence of conflict characteristics on corporate engagement was possible only because the researchers chose cases in regions with different conflict types and conflict intensity. For example, the violent conflicts examined in Nigeria were closely related to its oil production and the presence of oil companies there, whereas the sources of the Israeli–Palestinian conflict were not related to the activities of the tourism industry. It was assumed that such differences would be relevant in explaining corporate engagement.

Usually the process of case selection also includes pragmatic considerations. Advantages for a study include previous experience in certain countries or world regions and proficiency in the respective languages, which may heavily influence case selection. Access to the field is also a deciding factor. This issue presented itself in two variants in our research. First, access to the respective countries was required because field research was deemed necessary in analysing corporate engagement in conflict zones; however, in some cases the security situation or other travel restrictions presented difficulties (see Chapter 13 by Sabrina Karim). Second, access was needed to companies. Researchers contacted companies at an early stage with requests for further information and/or interviews (see below). The success of such requests greatly depended on who was approached at the company and on the previous experience that the person and the company in general had with participating in such research projects (see Chapter 5 by

Anja Mihr). Larger companies in particular might have in place general policies regarding how to deal with such requests. It is important to keep in mind that the first contact with responsible company staff might be more decisive than contacting public officials because the information to be obtained might be more widely distributed and therefore more easily available in the latter case. If interview requests were denied or ignored, a decision had to be made about whether a case study could be conducted without additional information from the company or whether another case study would be more suitable. This choice depended on the amount and quality of available academic literature,¹⁴ publicly available information, information available from corporate stakeholders and whether the case was central to the overall study. For example, analysing the role of the oil industry in Nigeria without covering the case of Shell would result in an incomplete picture, because the case is essential for understanding the broader picture (Zimmer 2010).

Methods ‘in use’

A combination of methods was used to obtain data on the output and outcome dimensions of corporate engagement in conflict zones. Based on the idea of triangulation, researchers used content analysis of documents, interviews and field research in conflict zones to generate data.

The term ‘triangulation’ refers to the ‘combination of two or more theories, data sources, methods, or investigators in the study of a single phenomenon’ (Kimchi *et al.* 1991: 384; see also Denzin 1970). Different types of triangulation are usually distinguished in the literature. Denzin (1970) distinguishes among four types: theory triangulation, investigator triangulation, data triangulation and methods triangulation. Theory triangulation involves the use of two or more (rival) theories to explain a given phenomenon, whereas investigator triangulation means that two or more researchers analyse the phenomenon in question. In data triangulation, multiple data sources are used to validate results. Methods triangulation can take two forms: within-method and between-method triangulation. Within-method triangulation involves the use of two or more approaches to collecting data for measuring a phenomenon. But all these approaches belong to either the qualitative or the quantitative ‘camp’. Between-method triangulation uses both quantitative and qualitative data collection approaches in one study to examine a particular variable (Denzin 1970; Flick 2011).

Our approach to studying the role of business in conflict zones can be related to the ideas of within-method and data triangulation as means of achieving more reliable results. Data collection approaches included document analysis, interviews and observations during field research in conflict zones (i.e. within-method triangulation). Moreover, data were generated to examine the phenomenon of corporate governance contributions at different points in time, at different places and based on information from different individuals. In fact, one might call the type of triangulation used in our research project as ‘soft triangulation’, because the intention was to complement and confirm our findings rather than to challenge them.¹⁵

Content analysis of documents

Work on each company case study usually started with a review of the academic and the 'grey' literature, in addition to researching and analysing publicly available documents. The most important source proved to be documents published by the corporations themselves. In most cases, the documents that were available and easily accessible on corporations' websites were codes of conduct and CSR or sustainability reports; otherwise, the availability of documents on corporate activities in conflict zones varied from case to case. Content analysis of these documents not only constituted the primary evidence for inferring corporate policies (self-commitments), it also provided preliminary insights into corporate activities (outcomes) related to self-commitments, because companies report on both their policies and the implementation of such policies, particularly their perceived achievements.

Being widely accepted as a method for analysing governmental policies, analyses of *corporate* documents are often contested (for a discussion, see Kollman 2008). Corporate documents are often regarded as pure public relations vehicles and 'cheap talk'. Nevertheless, in order to deduce a corporation's self-commitments, it is essential and perfectly appropriate to take official corporate statements seriously. Because corporate reports are usually available online for several years, this type of analysis also allowed us to trace changes in a corporation's policies and activities over time. Other, supplementary sources of information were considered to the extent possible, including reports published by (critical) civil society organizations.

A content analysis¹⁶ (see also Chapter 6 by Joakim Berndtsson in Section I of this volume) was then conducted, during which the documents were screened for corporate policies and activities. In addition, the content analysis was used to infer explanations for corporate engagement. Using the project's codebook as a point of departure, we found that some explanations for corporate engagement were easily identifiable in corporate documents, such as references to the importance of a company's reputation, changing societal expectations or membership in business initiatives.

Short profiles of the corporations, including basic information on the company, as well as on its outputs and outcomes, were developed as a result of the document analysis. Although in many cases creating these corporate profiles had already begun during the case selection process, they were developed further in an iterative process. For example, during the field research, the corporations and other stakeholders often provided additional documents.

Expert interviews

Expert interviews were a major source of information. The interviewees were experts in the sense that they were part of the field of action that constituted the empirical case (Meuser and Nagel 2005: 73). These experts were expected to have insider knowledge about corporate engagement that could not be tapped by

relying only on an analysis of the public corporate documents. The interviews were developed as semi-structured interviews (see Chapter 5 by Anja Mihr) with the aim of gaining further insights into corporate outputs and outcomes, as well as into the different factors explaining a company's engagement. The interviews were semi-structured (Gläser and Laudel 2004: 39), meaning that the researchers were not bound by the order of the questions or their exact wording. The questionnaires were adapted to the specific situation and interviewee.

Based on the idea of data triangulation, interviews were conducted with representatives of corporations and a variety of stakeholders, such as civil society organizations (NGOs), collective business and CSR initiatives, as well as representatives of governments, development agencies and international organizations. The selection of interviewees was case-specific and was often developed only as part of the field research in the conflict zones (see below). Interviewees were usually asked to recommend additional experts, who were then contacted to request interviews (snowball sampling).

It is evident that access to and interviews with corporate representatives were of special importance for analysing the different dimensions of effectiveness of corporate engagement. The researchers chose various strategies. In some cases, interview requests were first submitted to corporate headquarters, which are often based in the United States or Europe.¹⁷ Whenever possible, interviews were carried out in person, with telephone or Skype as the second-best alternative. Certainly this endeavour is not only time-intensive but also requires appropriate funding. Interviewees were asked to provide the names of additional respondents, especially corporate representatives in conflict zones and stakeholders who either cooperated with or challenged the corporation.

Gaining access to corporations required a certain level of endurance on behalf of the researchers (see also Monahan and Fisher 2015). In addition to sending written requests, researchers participated in industry and CSR conferences to establish contact with corporate representatives. However, it needs to be kept in mind that responsible corporate staff are usually very busy, often receiving many interview requests, and they have different previous experiences with researchers. Moreover, the issue of operations in conflict zones is often considered a very sensitive one, where transparency on the part of corporations might result in only more criticism. In fact, interview requests were not always successful, a frustrating experience that required pragmatic responses: some case studies were developed based on publicly available documents and interviews with corporate stakeholders; in other cases, granting anonymity was required to get access to a corporation (see Feil 2012); and as a last resort, researchers revisited the original case selection.

Field research in conflict zones

Field research in conflict zones was conducted to obtain additional data that could then be used to analyse corporate engagement on the ground. The main aim was to gain insights into the outcome dimension of effectiveness and to find

explanations for corporate engagement. Field research was regarded as an essential element of the research project and was deemed particularly important for studies analysing local small and medium-sized companies (see Fischer 2011). Publicly available information on such local companies (from corporate websites) is often not as rich as it is for transnational corporations. In addition, these local companies usually do not receive much attention from civil society and academic researchers, which is why in most cases information about these companies is more limited than is information about large corporations. The information obtained during field research was also important for studies focusing on transnational corporations to broaden and validate data gained through interviews at headquarters. This was achieved mainly through interviews with local company representatives and representatives of stakeholder groups. Other methods used during field research included observation during company visits and participation in academic and industry conferences.¹⁸ Even more important would be field research for gaining insights into the *impact* of corporate engagement on the local level. For example, analysing the impact of extractive operations often involves visits, interviews and focus group discussions with the local population affected by corporate operations and other stakeholders, which can be done only through field research.

Researchers usually arranged for one or two visits to ‘their’ respective conflict zones, with stays of three weeks to two months. The duration of a field stay depended on a variety of factors, such as the number of places the researcher wanted to visit and previous knowledge about the selected conflict zone, but also, to a degree, on personal preferences. Conducting field research involves specific practical, ethical and methodological challenges (see Chapter 12 by Tessa Diphoom and Chapter 13 by Sabrina Karim; Goodhand 2000; Wood 2006; Browne and Moffett 2014). In general, conducting some kind of field research has become *en vogue*, particularly among graduate students and junior researchers (Mitchell 2013; Browne and Moffett 2014). However, researchers must prepare carefully for such visits. Access to conflict zones is not always easy – due to visa restrictions or because of the local (security) situation. Field research often requires special permits or visas. It is equally important to be flexible and monitor the local (security) situation closely so plans can be adapted accordingly. In some cases, it might be advantageous to cooperate with local actors, such as civil society organizations, universities or development agencies, who can help in obtaining access or providing a workspace.

In addition to the many practicalities to be considered, ethical issues are also important. One important standard is to apply the generally accepted principle of ‘do no harm’, which means that researchers’ interventions must not cause any harm to participants in the research, such as interviewees (Goodhand 2000; Wood 2006). It is important to keep in mind that

[...] the ethical imperative of research (‘do no harm’) is intensified in conflict zones by political polarization, the presence of armed actors, the pre-

carious security of most residents, the general unpredictability of events, and the traumatization through violence of combatants and civilians alike.

(Wood 2006: 373)

Ethical problems are most likely to occur when researchers get in touch with the population and other stakeholders who are affected by corporate operations in conflict zones. In the context of business in zones of conflict, almost all the researchers in our project perceived the field experience as emotionally stressful at some point because of their first-hand insights into the local populations' living conditions, whether or not these conditions were related to corporate activities. Ethical considerations also played a role when critical civil society representatives were interviewed. Although an invaluable source of information, they might not dare to speak openly against corporate and/or state officials. In general, interviewees must understand the risks and benefits of their participation in a research project and must agree by giving informed consent (Wood 2006: 379–380). Practical issues involved in the planning of interviews also had to be considered, such as finding a time and place that was perceived by both sides as being safe. Another sensitive issue was how to ensure security of the data and anonymity of the interviewees. For example, names of interviewees should be stored separately from interview records and transcripts.¹⁹

Another aspect of field research is the possibility that interviewees might expect something in return for their participation, thus imposing (emotional) pressure on the researcher. Most often, the interviewees in our project simply expected their story to be told, but sometimes material or other benefits were expected. For example, this issue was broached during interviews with civil society representatives in Nigeria, who expected that the researchers would have some useful connections with development agencies or foundations in their home country that might provide funding for certain projects. Therefore, it is important for the researcher to clearly articulate the goal and status of the study and the researcher, for example by communicating that one is not affiliated with a certain corporation or development agency. At the same time, it is appropriate to reflect on what to give back to those who contributed to the research. The literature on field research mentions such options as disseminating publications (and translating them into local languages as necessary) or returning field notes and other materials to the individuals involved (see Chapter 12 by Tessa Diphooorn; see also Wood 2006).

Finally, when working on and with corporations, it is important for the researcher to maintain a neutral stance. Access to corporations is needed to gain access to information and company sites and, in some instances, to travel safely to remote places. Corporations might offer guided visits to production sites and/or communities. However, being accompanied by a corporate official or even using transportation provided by corporations might influence who is willing to talk and what answers the researcher might get. At that point, emphasizing that one is an independent researcher and does not work for the company might be of only limited value. If possible, an alternative may be to arrange visits with staff from local development agencies or civil society organizations.

In summary, it is absolutely necessary to reflect on the necessity of field research as well as on the methods and procedures to be used. Very practical advice in the context of preparing field research is to talk to experienced colleagues with in-depth knowledge about a certain country or region. They are usually willing to share their experiences and best practices.

Results

After we had applied the methods and the conceptualization of effectiveness described in this chapter, the project revealed a significant level of corporate governance engagement in conflict zones. However, ‘the most striking empirical finding [was] the virtual lack of any direct contributions to security by corporations in almost all case studies’ (Deitelhoff *et al.* 2010: 204). As was suggested by many corporate interviewees, this is mainly because security issues are often regarded as belonging to the core responsibility of the state. Engaging in security governance continues to be a taboo for companies. Corporate governance engagement (outputs and outcomes) that was identified was rather indirectly related to peace and security by focusing on issues such as development, combating corruption, environmental protection or human rights protection. However, there were considerable differences among companies, industries and conflict zones.

Another interesting finding was a general trend towards developing CSR and governance programmes over time, mainly since the 1990s. This could be explained by the significance of a particular background factor: almost all case studies identified the importance of a changing transnational normative environment that pushes corporations to make more governance contributions. In other words, these emerging expectations led companies to develop, revise and extend their CSR policies and related activities (Deitelhoff *et al.* 2010: 205–206). Therefore, although expectations regarding corporate engagement in conflict zones must not be too high, corporations are relevant players in conflict zones and should not be overlooked.

Conclusion

How effective is the engagement of corporations – and of non-state actors in general – in conflict zones? This is a pertinent research question for global governance and peace and conflict research, not least because evaluating such engagement will provide an important, though certainly not the only, argument for bringing non-state actors into governance arrangements. The lesson learnt from our research is that addressing this research question is feasible and important but requires careful methodological reflection and contextualization of the results. The conceptual and methodological toolbox developed in regime theory has been applied not only to interstate but also to public–private regimes, and it can also be adapted to analyse non-state actors’ interventions in conflict zones.

Focusing research on the output and outcome dimensions of effectiveness has advantages and disadvantages, as discussed in this chapter. One important contextualization of results is that researchers can usually examine only one or a few cases of corporate engagement in conflict zones owing to their limited capacities. Results cannot and should not be overstretched and usually do not refer to the business community as a whole to avoid the risk of holistic fallacy.

It is equally important to keep in mind that, despite the challenges, analysing corporate impact is an important scholarly task. The outcome and impact of corporations need to be kept as separate dimensions in scholarly and public debates. It is reasonable to assume that a gap exists not only between output and outcome – the argument being that corporate CSR is merely window dressing – but also between outcome and impact. This gap might be due to a variety of factors, including the fact that even well-intended corporate programmes may not have the desired impact and that a variety of factors influence the course of violent conflicts and lie beyond the corporate sphere of influence. To measure corporate impact one would require a different methodological toolbox. For example, intensive field research is usually needed for measuring corporate impact on communities. In such cases, researchers usually stay for extended periods of time in the communities where corporate activities are located in order to interview members of the local communities and/or to organize focus group discussions (see, for example, Jacobs 2015). Another important dimension of effectiveness is analysing unintended consequences, which may be positive or negative and occur on different levels. One important consequence of non-state actors' engagement in conflict zones may be their interaction with the responsibility of the state for providing basic public goods. Although non-state actors are usually brought in to compensate for some kind of state failure, their engagement may, in the long run, further weaken the role of the state rather than strengthen it.

Notes

- 1 The authors would like to thank the editors for their valuable comments on an earlier version of this chapter and Olga Perov for her research assistance.
- 2 The term 'governance' has become prominent in the fields of Political Science and International Relations since the 1990s. Broadly defined, governance involves all forms of collective regulation, including private self-regulation, different forms of co-regulation and (inter-)governmental regulation. The use of the term thereby represents a departure from the analytical focus on governments as the sole providers of public goods to also encompass other international and non-state actors who exercise authority (Rosenau 1995; Avant *et al.* 2010).
- 3 The project was funded by the Fritz Thyssen Foundation and the Cluster of Excellence 'Formation of Normative Orders'. Its main results were published in Deitelhoff and Wolf (2010b), Fischer (2011), Haidvogl (2011) and Feil (2012).
- 4 The research project focused on transnational and local companies that do business in zones of (violent) conflict. Actors who provide security as part of their core business, such as private military and security companies, were excluded from the analysis.
- 5 For example, it is obvious that corporate investments and business activities are linked to economic growth – a relationship that could be interpreted as an important

corporate contribution to the peaceful development of a society; however, mere business activities do not fulfil the requirement of intentionally contributing to governance. Moreover, following legal standards is compulsory for corporations. In some countries, for example, companies are required by law to pay certain social contributions to the state or communities; such legally mandated payments are therefore not voluntary governance contributions.

- 6 The information in this section is based on Wolf (2010).
- 7 See, for example, the contributions in Hegemann *et al.* (2013) and those related to regime theory: Young and Levy (1999) and Young (2004).
- 8 The distinction is also widely used in evaluation research, such as in studies that evaluate development projects.
- 9 One important difference is that research on regimes usually focuses on an existing international regime, the effects of which are to be evaluated (Underdal 1992). In contrast, when doing research on business in zones of conflict, the focus is on the interventions of a single actor.
- 10 In the context of the research project, this was coded as a pro-active commitment.
- 11 In intrastate conflicts, the level of violence depends heavily on the relationships among the different parties to the conflict (e.g. the state and some non-state armed group or groups). A corporation is not necessarily a party to the conflict, and its impact on the evolution of a conflict might well be limited. Corporate impacts might be easier to identify on the local level; for example, in conflicts involving large-scale extractive projects, corporations are often influential parties, because their policies and behaviour directly affect the level of violence.
- 12 Given the unsatisfactory state of the research, the approach chosen was rather an inductive one. Certain working conjectures about corporate engagement were being tested for their plausibility (see Eckstein 1975), but there was also scope to identify other important variables along the way.
- 13 We analysed data from different conflict databases, including the Uppsala Conflict Data Program (2015), the Arbeitsgemeinschaft Kriegsursachenforschung (2015) and the Conflict Barometer 2015 of the Heidelberg Institute for International Conflict Research (HIIK 2015). Although the focus of the project was on intrastate conflicts, these conflicts – particularly those in sub-Saharan Africa – are embedded in complexes of regional conflicts.
- 14 When one is conducting research on business, important sources of information are usually the academic CSR and business journals.
- 15 See Turner and Turner (2009) for the distinction between soft and hard triangulation.
- 16 The term ‘content analysis’ is used widely and includes quantitative and qualitative, as well as more inductive and deductive, forms (for an overview, see Mayring 2010; Krippendorff 2013).
- 17 Obviously, this strategy was possible only for research on transnational corporations, not for research on small and medium-sized corporations headquartered in conflict zones.
- 18 Note that our approach to field research differed from anthropological approaches, in which participant observation is a central component (see Chapter 12 by Tessa Diphooorn). A focus on participant observation would require much longer field trips and would be of particular use in answering research questions other than assessments of effectiveness.
- 19 Other ethical dilemmas may include how to present sensitive material in publications and how to thank those who supported and participated in the research (Wood 2006: 382).

References

- Arbeitsgemeinschaft Kriegsursachenforschung. 2015. *AKUF Database on Wars* (online). Available from: www.wiso.uni-hamburg.de/fachbereiche/sozialwissenschaften/forschung/akuf/akuf/ (accessed 10 December 2015).
- Avant, Deborah D., Martha Finnemore, and Susan K. Sell. 2010. 'Who Governs the Globe?', pp.1–34 in *Who Governs the Globe?*, D.D. Avant, M. Finnemore, and S.K. Sell, eds. Cambridge: Cambridge University Press.
- Beisheim, Marianne, and Andrea Liese, eds. 2014. *Transnational Partnerships: Effectively Providing for Sustainable Development?* Houndmills: Palgrave Macmillan.
- Browne, Brendan, and Luke Moffett. 2014. 'Finding Your Feet in the Field: Critical Reflection of Early Career Researchers on Field Research in Transitional Societies'. *Journal of Human Rights Practice* 6(2): 223–237.
- Coni-Zimmer, Melanie. 2014. *Corporate Social Responsibility zwischen globaler Diffusion und Lokalisierung: Britische und chinesische Ölindustrie im Vergleich*. Baden-Baden: Nomos.
- Dany, Charlotte. 2013. 'Between Big Deals and Small Steps: Measuring the Effectiveness of International Non-Governmental Organizations', pp. 81–99 in *Studying 'Effectiveness' in International Relations: A Guide for Students and Scholars*, H. Hegemann, R. Heller, and M. Kahl, eds. Opladen: Barbara Budrich.
- Deitelhoff, Nicole, and Klaus Dieter Wolf. 2010a. 'Corporate Security Responsibility: Corporate Governance Contributions to Peace and Security in Zones of Conflict', pp. 1–25 in *Corporate Security Responsibility? Corporate Governance Contributions to Peace and Security in Zones of Conflict*, N. Deitelhoff and K.D. Wolf, eds. Houndmills: Palgrave Macmillan.
- Deitelhoff, Nicole, and Klaus Dieter Wolf, eds. 2010b. *Corporate Security Responsibility? Corporate Governance Contributions to Peace and Security in Zones of Conflict*. Houndmills: Palgrave Macmillan.
- Deitelhoff, Nicole, Moira Feil, Susanne Fischer, Andreas Haidvogel, Klaus Dieter Wolf, and Melanie Zimmer. 2010. 'Business in Zones of Conflict and Global Security Governance: What Has Been Learnt and Where to from Here?', pp. 202–226 in *Corporate Security Responsibility? Corporate Governance Contributions to Peace and Security in Zones of Conflict*, N. Deitelhoff and K.D. Wolf, eds. Houndmills: Palgrave Macmillan.
- Denzin, Norman K. 1970. *The Research Act: A Theoretical Introduction to Sociological Methods*. Chicago: Aldine.
- Easton, David. 1965. *A Systems Analysis of Political Life*. New York: Wiley.
- Eckstein, Harry. 1975. 'Case Study and Theory in Political Science', pp. 79–137 in *Strategies of Inquiry*, F.I. Greenstein and N.W. Polsby, eds. Reading: Addison-Wesley.
- Feil, Moira. 2012. *Global Governance and Corporate Responsibility in Conflict Zones*. Houndmills: Palgrave Macmillan.
- Feil, Moira, Susanne Fischer, Andreas Haidvogel, and Melanie Zimmer. 2008. *Bad Guys, Good Guys, or Something in Between? Corporate Governance Contributions in Zones of Violent Conflict*. PRIF Report No. 84. Frankfurt am Main: Peace Research Institute Frankfurt.
- Fischer, Susanne. 2011. *Reisen für den Frieden? Engagement von Unternehmen der Tourismusbranche in Israel und den Palästinensischen Autonomiegebieten*. Baden-Baden: Nomos.
- Flick, Uwe. 2011. *Triangulation: Eine Einführung*. 3rd edn. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Flohr, Annegret, Lothar Rieth, Sandra Schwindenhammer, and Klaus Dieter Wolf. 2010. *The Role of Business in Global Governance: Corporations as Norm-Entrepreneurs*. Houndmills: Palgrave Macmillan.
- Freeman, Bennett, and Genoveva Hernández Uriz. 2003. 'Managing Risk and Building Trust: The Challenge of Implementing the Voluntary Principles on Security and Human Rights', pp. 243–259 in *Business and Human Rights: Dilemmas and Solutions*, R. Sullivan, ed. Sheffield: Greenleaf.
- George, Alexander L., and Andrew Bennett. 2004. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Gerring, John. 2008. 'Case Selection for Case-Study Analysis: Qualitative and Quantitative Techniques', pp. 645–684 in *The Oxford Handbook of Political Methodology*, J.M. Box-Steffensmeier, H.E. Brady, and D. Collier, eds. Oxford: Oxford University Press.
- Gläser, Jochen, and Grit Laudel. 2004. *Experteninterviews und qualitative Inhaltsanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Goodhand, Jonathan. 2000. 'Research in Conflict Zones: Ethics and Accountability'. *Forced Migration Review* 8: 12–15.
- Haidvogel, Andreas. 2011. *Unternehmen als Friedensmacher? Zur Rolle privatwirtschaftlicher Akteure im nordirischen Friedensprozess*. Baden-Baden: Nomos.
- Hauffler, Virginia. 2010. 'Corporations in Zones of Conflict: Issues, Actors, and Institutions', pp. 102–130 in *Who Governs the Globe?*, D.D. Avant, M. Finnemore, and S.K. Sell, eds. Cambridge: Cambridge University Press.
- Hegemann, Hendrik, Regina Heller, and Martin Kahl, eds. 2013. *Studying 'Effectiveness' in International Relations: A Guide for Students and Scholars*. Opladen: Barbara Budrich.
- HIK (Heidelberg Institute for International Conflict Research). 2015. *Conflict Barometer*. Available from: www.hik.de/en/konfliktbarometer (accessed 10 December 2015).
- Huckel, Carmen, Lothar Rieth, and Melanie Zimmer. 2007. 'Die Effektivität von Public–Private Partnerships', pp. 115–144 in *Macht und Ohnmacht internationaler Institutionen*, A. Hasenclever, K.D. Wolf, and M. Zürn, eds. Frankfurt am Main: Campus.
- Jacobs, Andreas. 2015. *Assessing Corporate Impacts: Responsible Mining and Local Development in Africa*, PhD Thesis, Department of History and Social Sciences, Technische Universität Darmstadt.
- Kimchi, Judith, Barbara Polivka, and Joanne S. Stevenson. 1991. 'Triangulation: Operational Definitions'. *Nursing Research* 40(6): 364–366.
- Kollman, Kelly. 2008. 'The Regulatory Power of Business Norms: A Call for a New Research Agenda'. *International Studies Review* 10(3): 397–419.
- Krasner, Stephen D., ed. 1983. *International Regimes*. Ithaca: Cornell University Press.
- Krippendorff, Klaus. 2013. *Content Analysis: An Introduction to Its Methodology*. 3rd edn. Los Angeles: Sage.
- Mayring, Philipp. 2010. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 11th ed. Weinheim: Beltz.
- Meuser, Michael, and Ulrike Nagel. 2005. 'ExpertInneninterviews – vielfach erprobt, wenig bedacht: Ein Beitrag zu qualitativen Methodendiskussion', pp. 71–93 in *Das Experteninterview: Theorie, Methode, Anwendung*, A. Bogner, B. Littig, and W. Menz., eds. 2nd edn. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mitchell, Audra. 2013. 'Escaping the "Field Trap": Exploitation and the Global Politics of Educational Fieldwork in "Conflict Zones"'. *Third World Quarterly* 34(7): 1247–1264.

- Monahan, Torin, and Jill A. Fisher. 2015. 'Strategies for Obtaining Access to Secretive or Guarded Organizations'. *Journal of Contemporary Ethnography* 44(6): 709–736.
- Reinicke, Wolfgang H., and Francis Deng. 2000. *Critical Choices: The United Nations, Networks, and the Future of Global Governance*. Ottawa: International Development Research Centre.
- Rosenau, James N. 1995. 'Governance in the Twenty-First Century'. *Global Governance* 1(1): 13–48.
- Turner, Phil, and Susan Turner. 2009. 'Triangulation in Practice'. *Virtual Reality* 13(3): 171–181.
- Underdal, Arild. 1992. 'The Concept of Regime "Effectiveness"'. *Cooperation and Conflict* 27(3): 227–240.
- Underdal, Arild. 2004. 'Methodological Challenges in the Study of Regime Effectiveness', pp. 27–48 in *Regime Consequences: Methodological Challenges and Research Strategies*, A. Underdal and O.R. Young, eds. Dordrecht: Kluwer Academic.
- Uppsala Conflict Data Program. 2015. *UCPD Conflict Encyclopedia* (online). Available from: www.pcr.uu.se/research/UCDP (accessed 10 December 2015).
- Vogel, David. 2005. *The Market for Virtue: The Potential and Limits of Corporate Social Responsibility*. Washington, DC: The Brookings Institution.
- Wenger, Andreas, and Daniel Möckli. 2003. *Conflict Prevention: The Untapped Potential of the Business Sector*. Boulder: Lynne Rienner.
- Wolf, Klaus Dieter. 2010. 'Output, Outcome, Impact: Focusing the Analytical Lens for Evaluating the Success of Corporate Contributions to Peace-Building and Conflict Prevention'. *PRIF Working Paper* 2. Available from: http://hsfk.de/fileadmin/downloads/PRIF_WP_03.pdf (accessed 31 December 2015).
- Wolf, Klaus Dieter, Nicole Deitelhoff, and Stefan Engert. 2007. 'Corporate Security Responsibility: Towards a Conceptual Framework for a Comparative Research Agenda'. *Cooperation and Conflict* 42(3): 294–320.
- Wood, Elisabeth Jean. 2006. 'The Ethical Challenges of Field Research in Conflict Zones'. *Qualitative Sociology* 29(3): 373–386.
- Young, Oran R. 2004. 'The Consequences of International Regimes', pp. 3–24 in *Regime Consequences: Methodological Challenges and Research Strategies*, A. Underdal and O.R. Young, eds. Dordrecht: Kluwer Academic.
- Young, Oran R., and Marc A. Levy. 1999. 'The Effectiveness of International Environmental Regimes', pp. 1–32 in *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioural Mechanisms*, O.R. Young, ed. Cambridge, MA: MIT Press.
- Zimmer, Melanie. 2010. 'Oil Companies in Nigeria: Emerging Good Practice or Still Fuelling Conflict?', pp. 58–84 in *Corporate Security Responsibility? Corporate Governance Contributions to Peace and Security in Zones of Conflict*, N. Deitelhoff and K.D. Wolf, eds. Houndmills: Palgrave Macmillan.

15 Discussion chapter

Practising reflexivity in field research

Jacqui True

In this commentary, I reflect on three distinct approaches to field research for studying distinct actors in the arena of international security: a field experiment designed to test the impact of gender reforms within policing in a single site, post-conflict Liberia, a comparative case study analysis assessing the effectiveness of business actors' corporate social responsibility initiatives to address community security in a range of sites, and an ethnographic study of private security officers in several conflict-affected cities. All approaches come with their strengths and their limitations, and these three studies bring fascinating insights into new and emerging actors in the security field. We may have the best ideas, research questions and hypotheses in the world but the rubber really meets the road when we consider how to study them empirically. We must be able to answer our questions through a method or process that maximizes the chances that our answers are meaningful and defensible, withstanding critical scrutiny and alternative findings and explanations. I would argue that being able to devise a methodological approach that includes a research question, research design as well as specific methods for collecting, analysing and disseminating data and knowledge is not just a practical problem for the researcher; it is also a profoundly ethical and theoretical challenge.

The authors of the chapters in this volume have been asked to consider the practical application of their respective methods, including – in this section – field research methods. They have been encouraged to consider which criteria are important in selecting a suitable method for a particular research question, how a method (or a combination of methods) can be translated into a research design, and how to resolve practical issues with the application of a method to a particular case. They have also been asked to reflect on the relative interpretive or explanatory power of a particular analytical method and the sort of knowledge that is gained and the insights that may be neglected when a certain method is used. These common questions about the implications of different methods allow us to compare various approaches in a common field. However, they don't allow us to interrogate where a researcher's questions come from in the first place, which likely influences the choice of methods, nor do they give us an account of the researcher's 'situatedness' in the research field.

As researchers, we are part of the social world and not separate from the field we study. Our choices of study and of methods are therefore inherently social

and normative; that is, they have an effect on the world. Social science informed by positivist epistemology rests on the stance of objectivity, which assumes that the researchers can remove the effects of their subjectivity in the research process by following a replicable method. In describing their field research methods, each of the three chapters in this section recounts the authors' efforts to achieve 'objectivity'. However, the quest for objectivity itself may end up masking important social and political dynamics, including the situatedness of the researcher in the research subject/object. Our study of a subject may produce changes in that subject in ways that we will either neglect and/or not be adequately aware of unless we are reflexive. This notion of inter-subjectivity is recognized in the physical sciences as the 'observer effect', where the act of measuring certain systems is known to create change in the systems themselves.

To address the intersubjective dimension of research, social science informed by post-positivist epistemologies seeks to give a full account of the research process as a social and ethical process (Jackson 2011). Critical feminist methodology, for instance, builds reflexivity into all stages of the research process as a core ethic that improves our research (Ackerly and True 2008, 2010). It aims to be reflexive about the windows and the blind spots opened up by all methodological choices from the choice of research question and design through to the specific methods and strategies used to collect and analyse data. Feminist methodology seeks to achieve 'strong objectivity' precisely by socially situating that knowledge and being conscious of the effects of subjectivities – and not just in terms of their potential bias and limitations but rather, in terms of the power of certain 'standpoints' and situatedness of researchers to generate critical knowledge (Harding 1991). Critical knowledge is knowledge that is able to scrutinize the origins, sources and interests associated with dominant forms of knowledge to gain a better and more complete purchase on the social world, be it the world of non-state actors in international security or any other world.¹

Critical feminism involves commitments (1) to being attentive to the power of different epistemologies in the research process; (2) to boundaries that include/exclude and often marginalize; (3) to all relationships in the research process including that between the researcher and the researched; and (4) to situating oneself as a researcher in the field. When we critically reflect on these four dimensions, power is the researcher's subject rather than the researcher being the agent of power (Ackerly and True 2008: 699).

In the remainder of this chapter, I explore the opportunities and challenges of the research question and methodological approach in each of the three field research studies and how each practises reflexivity within and about its methods and methodology.

Field experiments for studying gender and security sector reform

Sabrina Karim's study examines the effects of gender reform in the Liberian National Police – a post-conflict, if not a non-state, actor. She asks, how does the

expansion of these reforms affect civilian perceptions and support for government? The two major reforms under study appear to be the integration of women into the police as well as training to ensure police report sexual and gender-based violence (SGBV) incidents within neighbourhoods. The prevalence of SGBV after official conflict has ended is a significant issue in many post-conflict contexts. Policymakers in the United Nations and other international organizations and donor states have advocated for increasing the number of women employed in the security sector as one way to improve the responsiveness of governments and security sector to this problem of violence. However, the solution – despite generating considerable political support and some financial commitments – is largely untested.

In this study, Karim wanted to find out whether the gender of police officers makes a difference to the reporting of SGBV crimes. She and her colleagues designed a study as close as possible to the randomized control trials associated with medical trials – the high water mark for ‘science’. Her theoretical approach, which involves an isolated variable of interest – the gender of officers in patrols – and a seemingly parsimonious hypothesis, emanating from the UN, to test appears to match the design requirements for field experiments. In the Liberian police field experiment, Karim and her collaborators used experimental conditions to control for the presence and absence of female sex/gender of the officer patrol, of police visits to neighbourhoods, and of prior information shared by police in neighbourhoods. Experimental methods aim to mitigate bias or the effects of subjectivity such as endogeneity, omitted variable and selection bias often associated with observation methods, especially through statistical inference.

However, these issues may also inhere, even in experimental methods or other methods explicitly designed to eliminate entirely the effects of subjectivity and the larger social world from the particular subject or unit of analysis under study. For example, the problem of endogeneity is possible. For instance, in the experiment conducted by Karim and her colleagues we cannot ignore the possibility that the outcomes of women’s inclusion in the Liberian police force is related to the outcomes of SGBV reporting to the extent that both are likely causal effects of the broader situation of women’s empowerment in Liberia. Researchers also have to ask about the effects that the selection of a particular country or region might have on the outcomes of the research. In the current case, the selection of Liberia as the site for the field test may also import selection bias. Liberia is not an average post-conflict country when it comes to gender reform. Indeed, the female President made gender equality a key platform of her government and explicitly requested an all-female UN police support unit to protect her government (Pruitt 2016). Thus, the Liberian police force may not fulfil all the specifications of scope conditions to enable it to serve as a generalizable case. The findings for Liberia may be derivative of other, prior overarching state-building reforms not tested for in the study on the effects of gender reforms. This potential raises the possibility of another form of bias associated with omitted or missing variables not controlled for in the research design.

Hence, researchers should always have in mind that they might not be able to control for or include all possible variables in their research design.

No research design can ever approximate the complexity of the social world by operationalizing and randomizing all the factors affecting outcomes, such as change in reporting behaviour. Equally, no method in the social world can easily isolate the effects of one or two factors – the sex/gender of police officers and neighbourhood presence – over others, such as the other personal identity attributes of police, small group dynamics, prior education and so on. Nonetheless, a field experiment makes sense as a costly but feasible way to provide evidence testing a limited but powerful, singular theory that women police officers or gender balance among police increase citizens' – especially women citizens' who are likely the majority of victims of SGBV – sense of trust and security in government. For more sophisticated research questions that seek to understand the reporting of SGBV crimes and all the factors affecting this reporting or lack of it, we would need a different method and likely one that involves qualitative research in particular contexts (Davies *et al.* 2016). One-off experiments, moreover, can only assess short-term outcomes and may not be able to capture changes in reporting practices over a longer time period, which is typically needed to bring about, as well as ascertain, social and normative changes. Hence, while a particular (experimental or non-experimental) research design might provide useful insights on a particular issue, researchers always have to be aware of the limits of their research.

Reflexivity in field experiments

Sabrina Karim highlights how a reflexive approach to the field experiment method was crucial to its successful application. Such an approach goes beyond the requirements of institutional 'human subject ethics approval' processes within universities and other research organizations and the 'do no harm' principle common in other research. She discusses the power of researchers to decide which site or community receives an intervention or treatment and which serves as the control group and the effects this power may have on people's lives. Finding ways to mitigate this power and the unfairness of its impact are crucial considerations for researchers using the field experiment method. For instance, it may involve negotiating with authorities to enable the intervention/treatment to be rolled out to control groups at the end of the study. Karim also highlights the power relations between researchers and local enumerators who are often involved in recording results during the field experiment. As a way of mitigating a potentially extractive and exploitative relationship she suggests that researchers could contribute to skills development in the post-conflict country with training workshops and the like. Karim also encourages us to consider what knowledge our field research leaves behind and how we can give back to the communities engaged in our research as participation subjects. These are also key elements of a feminist research ethic that is reflexive about the effects of our research process as well as the impact of our research findings.

Case study analysis and policy effectiveness

The design and selection of cases to assess the effectiveness of security initiatives by business organizations in conflict-affected states requires the same rigour and involves similar challenges to the design and selection of a security sector field experiment site. In their chapter, Melanie Coni-Zimmer and Klaus Dieter Wolf present a qualitative case study approach to understanding the role of multinational business actors, their contributions to security initiatives and what effects they have. Given the powerful nature of the actors and the public scrutiny in the past over human rights abuses that have occurred during their business activities in high-risk settings, Coni-Zimmer and Wolf knew from the outset that gaining access to the subjects of this study was going to be difficult. (We gather also from Karim that it was time-consuming and difficult to negotiate access to the Liberian police force in order to be able to conduct a field experiment.)

Given the access issues but also the dangerous contexts and the language barriers in some of them, a perfect research design, including a range of businesses across regions and industries, was not practically feasible. The authors acknowledge that in the end the selection of business actors, although originally intended to address different regional conflict situations and types of businesses, was based on practicality: access, availability, data richness and language competency. This is so often the case – our methods must be fit for studying the so-called real world but they must be adapted for the real-world constraints on research. Moreover, because the study was dependent on access and information from the businesses themselves, the researchers found it challenging to maintain their impartiality. This is a common challenge for researchers when ‘studying-up’ among political and economic elites, as is often the focus of political science and International Relations studies not only in the realm of international security.

A key conceptual issue for the researchers in this study was being able to define and distinguish the outputs from the outcomes and the impacts of corporate security initiatives in the community. This is a familiar issue in policy research. Coni-Zimmer and Wolf note that while it is relatively easy, it is less theoretically interesting to measure outputs, such as the existence of policies and programmes as reflected in statements, codes of conduct and triple bottom line reporting, compared with outcomes such as behavioural and organizational change. To study outcomes we must look beyond tangible and objective things that we can read and count, to engage in qualitative research on intangible attitudes and practices. However, the challenge here, as with the field experiment, is to define outcomes in the most tangible way – such as a change in community leadership, in surveyed community opinions or in a reporting pattern of complaints and abuses, for example. Impact can be even more difficult to ascertain because the time line may be longer and causal chain(s) more challenging to document, and it may be difficult to isolate the contributions of other inputs. Here, qualitative analysis is most needed to be able to document and trace the

causal processes through which business initiatives led to a change in outcomes, as well as to contextualize these processes. While social scientists often put great stock on statistical correlations that reveal the strength of a relationship between an initiative and an outcome, they do not actually show the mechanisms that connect cause and effect. That requires at least some attempt at case analysis and counter-factual argument – ideally, case analysis where factors that represent counter-arguments and explanations are also present.

Comparative analysis of cases is also a good way to test the effects of certain factors that are present to different degrees in real world cases, e.g. the size and type of business actor, the level of investment in security programming, the design of the policy/programme, the conflict context, and so on. To enable analysis of the findings across the cases, Coni-Zimmer and Wolf produced a detailed codebook that specified the security initiatives under study and their various attributes as well as a range of explanatory variables that could affect the security initiatives and situation in positive or negative ways. With that explicit guide, the researchers could be confident that even in a study of a relatively small number of business actors they would isolate the key factors contributing to the effectiveness of security governance initiatives. They could also be confident that all researchers in the team were working with similar concepts and the one design.

Reflexivity in case study field research

The authors draw on a range of social science techniques to increase the significance of their findings for broader knowledge about business actors in security governance. They discuss how they ‘triangulated’ across multiple sources of data and methods for collecting the data in seeking to complement and confirm or falsify them. They mention the efforts made to ensure a diverse sample of corporations and, within each case study, a diverse sample of stakeholders through the snowball method. However, the anti-snowball method can also be useful in gaining a comprehensive sample. Including those who are least likely to be suggested by the corporations themselves is one way to triangulate the information of other stakeholders and to potentially falsify as well as confirm findings of effectiveness (Ackerly and True 2010). The authors do anticipate that there may be unintended consequences of business security initiatives and that they must be alert to looking for these and studying them. Reflexivity encourages us to question our expected answers and to anticipate the unanticipated in the process as well as the findings of research. As Cynthia Enloe (2011) says in *The Curious Feminist*: ‘If nothing ever surprises you then there would not be anything more to learn.’

With respect to reflexivity that addresses the impact of the researcher and the research process, Coni-Zimmer and Wolf are aware of the ethical issues and risks in working with local groups who may experience difficult living conditions, in part as a result of corporate initiatives. They discuss the risks and benefits to participant subjects of giving informed consent, the fact that researchers

may be seen as affiliated with the business organization rather than independent scholars. They mention the importance of sharing their data and analysis with participants in the form of translated publications and even field notes, which may be more accessible than academic writing. All these practices are crucial ways to ensure that our research does not harm and that we improve the world, not only with the addition of our research findings but also through our research process. That process attends to the relationships of power that are part of research, and we take responsibility for rebalancing these relationships in some way.

Participant observation and private security

Tessa Diphorn's chapter describes her approach to participant observation as a method for studying private security officers in South Africa. Her purpose was to understand how these security officers practice security, and she developed a new concept – 'twilight policing' – that conceptualizes the nature of the relationship between state and non-state providers of security as a result of her field research. Specifically, Diphorn analysed the routines of the private security officers in order to gain deeper insight into the meaning of the work and its interface with traditional public security organizations. Tacit routines are not readily viewable through one-off interviews, surveys or other data points. They are revealed only over time when, through critical engagement, the researcher is able to peel back the formal rules from the informal, non-verbalized but shared norms. Participant observation is somewhat of an oxymoron – if one is a participant, one can hardly be an observer. Participant observation is also contradictory in positing that the outsider provides critical knowledge; but being an outsider, by its nature, prevents full 'participant' immersion.

The history of this method belies an attempt at objectivity – distancing the researcher and their subjectivity even while immersed in the social world of the study's participants. Diphorn describes seven aspects of the participant observation method, which she treats as a synonym for ethnography: participation, open interviews, observation, field notes, longitudinal research, rapport and reflexivity. While her chapter is the only one on field research methods in this volume to explicitly discuss reflexivity, it is not until quite late in the chapter, for instance, that we learn she is a female researcher observing and accompanying mostly male private security officers into situations that include incidence of violence. We do not learn anything about her race/ethnicity positioning in the South African context and how that affected her research interactions. She notes succinctly that being a woman 'influenced the performance of masculine behaviour by armed response officers'. However, how and why that is the case is left unanalysed.

This situatedness of the researcher, including the race and gender differences between the researcher and her participant subjects, is crying out for a reflexive analysis. Diphorn does acknowledge the danger involved in her research and the 'emotionality of participation', but she hardly mentions the gender issues;

and yet because of the gendered nature of security work, employing largely male workforces, and the masculine qualities expected to be a feature of this work, the researcher's gender would not have been ignored by the participants in her study (Kumra *et al.* 2014). In fact, one can imagine many of her findings based on social interactions would have been enabled and indeed filtered through gender and race relations. It is hard (albeit possible) given existing gender norms across cultures, to envision a male researcher sharing emotions with his male security officer subjects to the same extent or in the same way (Parashar 2011).

In contrast to the other two chapters on field research methods, Diphoorn does not discuss precisely how she selected *which* private security officers to observe. Perhaps this is because the ethnographic method is considered valid in its own right and intended to generate particular meanings only obtainable through deep immersion rather than generalizable findings. In that sense, the method reflects a different epistemology compared with the field experiment and case study approaches. However, in a second study, Diphoorn studied public-private security assemblages in five different cities through the ethnographic method, although not involving full immersion, in order to produce broader comparative findings.

In her application of the ethnographic method, Diphoorn describes listening as well as participating in everyday conversations as opposed to conducting formal or semi-structured interviews with particular questions to address. Listening is a key method for reflexive approaches to research – as we cannot literally 'see' meaning but we may be able to hear it (Bickford 1996). She argues that small talk is especially relevant for discussing sensitive issues. It enables participants to disclose information that they would be unlikely to share in a more formal setting, where the researcher is relatively unknown to the subject. Diphoorn notes the importance of observing the surroundings, the seating and use of space in which conversations take place as well as the body language and other non-verbal signs that may become observable only over time. These factors help to reveal how officers *feel* as well as what they *say* and *do* about carrying out 'security' work in frequently violent and dangerous situations.

While the method of participant observation expands 'the field' of study both with respect to the time frame of observation and the depth of interactions, a key challenge is how to determine when you have observed enough. With the grounded theory approach, the researcher knows when to withdraw from the field because a saturation point has been reached with the data, which is the same with successive interviews, interactions, documents or other data collected (Charmaz 2006). Grounded theory therefore seems a crucial complementary approach for the ethnographic method – especially truncated versions of the method typically used outside of the anthropology discipline.

Conclusion

The three chapters on the distinct challenges and the opportunities inherent in field research each justify the validity of the knowledge produced through their

distinct field experiments, case studies and ethnographic methods. The authors respectively go beyond traditional 'scientific' approaches to methodological rigour by considering the ethical dimensions of field research. Moreover, they demonstrate reflexivity about the research process and how it may shape the research findings that result from their application of method. They understand that social science knowledge is not only knowledge for its own sake but that it may do harm and/or it may enable social and political change.

Note

- 1 Feminist standpoint theory argues, for instance, that marginalized groups are socially situated in ways that make it more possible for them to be aware of things and ask questions than it is for the non-marginalized groups; and, moreover, that research focused on power relations (read: politics) should begin with the lives of the marginalized.

References

- Ackerly, Brooke A., and Jacqui True. 2008. 'Reflexivity in Practice: Power and Ethics in Feminist Research on International Relations'. *International Studies Review* 10(4): 581–600.
- Ackerly, Brooke A., and Jacqui True. 2010. *Doing Feminist Research in Political and Social Sciences*. New York: Palgrave.
- Bickford, Susan. 1996. *The Dissonance of Democracy: Listening, Conflict, and Citizenship*. Ithaca, NY: Cornell University Press.
- Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: Sage.
- Davies, Sara E., Jacqui True, and Maria Tanyag. 2016. 'How Women's Silence Secures The Peace: Analysing Sexual And Gender-Based Violence in a Low Intensity Conflict'. *Gender and Development* 24(3): 459–473.
- Enloe, Cynthia. 2011. *The Curious Feminist: Searching for Women in a New Age of Empire*. Berkeley, CA: University of California Press.
- Harding, Sandra. 1991. *Whose Science, Whose Knowledge? Thinking from Women's Lives*. Ithaca, NY: Cornell University Press.
- Jackson, Patrick T. 2011. *The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the Study of World Politics*. Abingdon: Routledge.
- Kumra, Savita, Ruth Simpson, and Ronald J. Burke, eds. 2014. *The Oxford Handbook of Gender in Organizations*. Oxford: Oxford University Press.
- Parashar, Swati. 2011. 'Embodied "Otherness" and Negotiations of Difference: A Critical Self Reflection on the Politics of Emotions in Researching Militant Women'. *International Studies Review* 13(4): 687–708.
- Pruitt, Lesley. 2016. *The Women in Blue Helmets: Gender, Policing and the UN's First All-Female Peacekeeping Unit*. Berkeley, CA: University of California Press.

Conclusions and perspectives for researching non-state actors in international security



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

16 From cookbooks to encyclopaedias in the making

Methodological perspectives for research of non-state actors and processes

Anna Leander

The contributions to this volume helpfully highlight the methodological diversity, and also some of the specific difficulties, involved in researching Non-State Actors and Processes in International Security (NAPIS¹). In this concluding chapter, I wish to further underscore and contextualize what I read as the core insight emerging from the volume taken as a whole: that researching NAPIS demands an approach to method that is well informed, but also open, imaginative and ready to make unconventional methodological moves and combinations, as is this volume itself. I will insist that, as the contributions show, there are deeply ingrained reasons why narrow and restrictive approaches to method will do more to hamper than help NAPIS researchers. Some of these reasons are general and pertain to all research activities. But many of them are specific to NAPIS research. They are related to the complex relation between the state-centrism of the categories through which knowledge (of researchers as well as of the researched) is structured, and to the possibility of acquiring knowledge that defies these categories. Indeed, failure to problematize these categories and to incorporate them as objects of analysis (as opposed to the point of departure for analysis) will do more to distort, obscure or even obliterate NAPIS than to further explanations or understandings of them.

For this reason, it is of essence to abandon an approach that sees methods as recipes to be followed – a cookbook approach – and instead adopt an approach that is well informed, imaginative, open and evolving. I think of this latter approach as having much more in common with an online interactive encyclopaedia, such as Wikipedia, than with a cookbook. It has no fixed table of contents or structure, as entries can be added. It is constantly evolving and built through researchers' intuitions, experience, and entries. It is methods in the making, developed on the go, as it were.² Indeed, I would contend that such an approach to methods is the only viable methodological perspective for NAPIS research, as well as something NAPIS research can contribute to methodological perspectives more broadly. Herein lies the significance of this volume. I develop the argument in two steps. First, I argue that an 'encyclopaedias in the making' approach to methods matters generally, but especially for NAPIS research. I then

proceed to develop the argument by looking more closely at its significance for answering, but also for asking, questions about NAPIS. In the concluding section, I return to the broader implications of this argument for the methodological perspectives of NAPIS research and beyond.

Mastering the making of method matters for NAPIS research

It's not news that methods matter for research. Indeed, the publication of this volume confirms this truism. I nonetheless think it worthwhile to insist that the methods that matter are of the 'encyclopaedias in the making' kind, and that overall developments have made this kind of methods matter ever more for all research, and finally that it is particularly important for NAPIS research. Method, understood as the questions about *how* research is carried out and hence also *how* its findings are arrived at, stands at the heart of any research endeavour. Methods provide insight into the foundations of claims made in the research, the processes through which they were formulated and thus also their standing and relation to other kinds of claims. The *how* questions of methods provide essential information about the foundations of science. Not surprisingly, there is consequently a long tradition of debating what kinds of method are most adequate for different kinds of questions, which is precisely what most books on methods are preoccupied with.

More centrally for the argument in this conclusion, there is also a very long tradition of cautioning against methodological closure, as well as against the illusion that strict application of some specific method could produce definitive and sufficient knowledge. Sheldon Wolin, for example, defended the work of the 'Theorist' against the encroachment of the 'Methodists' who, according to him, restricted 'the 'reach' of theory by dwelling on facts that are selected by what are assumed to be the functional requisites of the existing paradigm', hence making it impossible to see that 'because facts are richer than theories, it is the task of the theoretical imagination to restate new possibilities' (Wolin 1969: 1082). Similarly, Hans-Georg Gadamer concludes his magnum opus *Truth and Method* by recalling that 'our entire exploration has shown that the security offered by the use of scientific methods is insufficient to guarantee truth.... What the methodological tool cannot offer, must be achieved through a discipline of questioning and researching' (Gadamer 1990 [1960]: 494 my translation). Hans-Georg Gadamer and Sheldon Wolin are resisting a narrow understanding of scientific methods that stifles research and knowledge generally, and in particular the knowledge they were interested in: political philosophy and hermeneutics respectively. This does not mean that they are opposed to discussions of *how* research is carried out. They both write extensively on the topic and have found it crucial. However, it does mean that they dissociate themselves from 'method' understood as a narrow, restrictive, and stifling imposition, and underscore the significance of a more open and imaginative approach to the *how* questions of research.

The caution against method as closure has lost none of its centrality. However, methods are now so indisputably central that the option Sheldon

Wolin and Hans-Georg Gadamer prefer – which is to distance themselves from methods, because they are so often narrowly understood – has been all but eliminated. Instead, the caution has taken the form of redefinitions of ‘method’ designed to broaden method to encompass more open and evolving ways of dealing with the *how* question. These redefinitions have been influential for three reasons. First, as the authoritative status of science (in universal and imposing singular) has given place to an understanding of science as a contested and ultimately profoundly uncertain activity, an open and broader approach to method has become more central. While method talk cannot solve the conundrums associated with the need of relating to and comparing different contradictory and often incompatible alternatives, it is a way of clarifying what alternative was selected and how it was construed (Abbott 2004: especially Chapters 1 and 7; Nowtny 2016). It is a short cut of sorts, as Andreas Armbrorst (see Chapter 2, this volume) argues when discussing definitions of terrorism. Second, as uncertainty penetrates the research process itself, the capacity to show that one has an answer to *how* to accommodate uncertainty, and therefore discover something novel, moves to the centre. This further consolidates the centrality of methods for the confirmation of scientific authority and validity (Law 2004). Method has to be redefined so as to leave space for ‘luck’ (see Chapter 6). Finally, and perhaps most significantly, research is dependent to an ever-increasing extent on competitive funding. In the distribution of this funding, ‘methods’ is a standard requirement (Przeworski and Salomon 1995; Mirowski 2011). The consequence of these developments is that the methodological literacy required to formulate a ‘methods section’ for a research project has become a fundamental requirement for any researcher and for the credibility of research. Arguably, methodological language has become a *sine qua non* for the bulk of contemporary research.

Not surprisingly therefore, treatises promoting methodological vocabularies have been burgeoning. This is true in the hard sciences (e.g. Stengers 1995: 102–113 *et passim*, or Feyerabend 2010), but it is perhaps most noticeable in research traditions that previously lacked such vocabularies, either because they dealt with the *how* of research in other terms, as did International Law or Critical International Relations (e.g. Salter and Mutlu 2012; Shapiro 2012; Aradau and Huysmans 2014) or because they are recently established and linked to technological developments (see for example Ruppert *et al.* 2013; Flam and Kleres 2015 and Chapter 10 by Alexander De Juan in this volume; Ulmer and Koro-Ljungberg 2015). NAPIS researchers are participating in this redefining effort, as the publication of this volume testifies. Overall, the consequence is that the narrow (cookbook) understanding of ‘methods’ is increasingly displaced by an open and malleable understanding of method and a growing readiness to make the evolving ‘social life of methods’ an object of study in its own right (Steinmetz 2005; Savage 2013; Boltanski 2014).

These general reasons for adopting an ‘encyclopaedias in the making’ approach to methods are compounded in the context of NAPIS studies, as the contributions in this volume show. Indeed, the specific difficulty involved in studying Non-State Actors and Processes in International Security, is that NAPIS

cross the categorical boundaries that usually serve to organize knowledge. They span the inside/outside, the public/private, the civilian/military and the peace/war divides that structure not only most theories, but also legal systems, policy-making, the institutional set up of national and international politics, the organizational structures of armed forces and companies and the practices not only of researchers but also of most actors.

These categories are practical and performative (Bourdieu 2003; Leander 2015–2016). Their significance, violence and performativity can hardly be overstated (Elshtain 1981; Walker 1988; Tronto 2013). In fact, because they are so powerful, these divides are not only constantly mobilized both by observers and observed, they more often than not make it all but impossible to see (or capture) the place of NAPIS. The NAPIS are seemingly obliterated by the divides. They are split up to neatly fit the categories and hence become elusive or disappear from view altogether. The analogy with the chimera is a useful one in this context (Graz 2008; Leander 2014; Hurt and Lipschutz 2016). This mythological fire-breathing monster with characteristics of the lion, snake and goat, gains its power from being an impossible creature that is therefore also elusive to the observer. If it is cut up and analysed as only goat or lion or snake, its significance is lost. The same is true of NAPIS. If they are cut up to be studied through the divides that usually organize knowledge and the theories and methods associated with them, their significance is lost.

Yet this is exactly what often happens. While it has become commonplace to include reference to the elusive line separating NAPIS from their state counterparts (see also Chapter 1 by Andreas Kruck and Andrea Schneiker and Chapter 14 by Melanie Coni-Zimmer and Klaus Dieter Wolf), analyses more often than not proceed as if this did not matter. The reason is straightforward. It is difficult to circumvent the great categorical divides and their related ‘methodological nationalism’ (Scholte 2005). The state has imposed its own view on the world. It is structuring vision and division and its categorizations have come to dominate. The state has a historically developed and deeply anchored monopoly on ‘symbolic violence’ which fashions not only social practices, but also theory, method and our innermost feelings and thoughts as individuals. The state therefore poses ‘an epistemological problem’ by structuring categories and thinking (Bourdieu 1994: 98–102, 2012: *et passim*). This problem is particularly acute for the NAPIS that tend to disappear when these categories are left unproblematized and their performative effects left unnoticed.

The making of open and imaginative methods therefore matters in general, but especially for NAPIS research, as this volume and all the contributions in it testify. I have just suggested that for NAPIS research, clarifying ‘the *how*’ (or the method) of research in a grounded and informed manner while keeping it open enough to actually make discoveries is not merely a matter of useful reflexivity, scientific rigour or funding requirements. It is essential for making the object of research visible, for capturing it, for circumventing the risk of distorting, obscuring and obliterating it altogether. The introduction uses the figure of a researcher whose methodological toolkit contains only a hammer, and who

therefore sees the world as full of nails, as an entry point for underlining the importance of the methodological imagination. The trouble for NAPIS research is that most methods are deeply fashioned by the great categorical divides that make the state an 'epistemological problem' and therefore tend to see nails of very many kinds but not of the NAPIS kind. When they are applied to hammer away, they hit at random and often damage the understanding of NAPIS.

Methods for answering and *asking* questions about NAPIS

The volume is organized around the assumption that the *how* of research is something that one begins to think about once one has indeed come to terms with what the question to be researched is, and that there should be consistency between what the question is and how it is investigated. This indeed is a sound, and conventional, approach to methods. However, as transpires from many of the contributions, the state as an 'epistemological problem' complicates the neat sequencing. The categorical divides fashion data collection, interpretation and the writing up so that these often have to be altered along the way in order to actually work. They also often impose changes on the question itself. There is a constant tweaking and transforming of the methods, but also a reformulation of the relationship between method and question where the openness of method becomes a way of re-formulating the research question. Researchers are, in other words, contributing to, transforming and construing not only the methods appropriate to studying NAPIS, but also the questions asked about them.

Making methodological openness and imagination central is perhaps not something the contributors to this volume would necessarily opt for if they had a choice (I don't know). The point is that because of the specific difficulties involved in studying NAPIS they have little choice. As the introduction points out, this is not necessarily because it involves secretive and secluded security matters. Of course it also does that. However, many NAPIS (including those hybrids that span the divides) have to advertise themselves. They put up web-pages. They participate in trade fairs and professional associations. They publish professional magazines. NAPIS personalities have published memoirs. Rebel groups, NGOs specialized in security and/or commercial security companies are willing to talk to researchers, who sometimes are embedded in them. The material researchers obtain in this way is of course and obviously marked by efforts to convey and promote specific images of NAPIS (see Chapter 14 by Melanie Coni-Zimmer and Klaus Dieter Wolf and Chapter 7 by Jutta Joachim). This said, it provides considerable insights into NAPIS that often appear easier to access than public security actors and processes (see Chapter 5 by Anja Mihr, Chapter 4 by Magnus Dau and Marlen Martin, and Chapter 6 by Joakim Berndtsson) or for that matter other market actors (Czarniawska 2016).

The difficulty particular to data collection about NAPIS is that the data more often than not is fashioned by the great divides and hence mostly inadequate for the purpose. As discussed by Alexander De Juan (Chapter 10), Sabrina Karim (Chapter 13) and Patrick Mello (Chapter 9), quantitative data is often inadequate

for serious NAPIS research because the kind of data needed spans the usual boundaries separating data categories (national/international, public/private, civilian/military) and must break with standard assumptions about the state. This cannot be simply remedied by collecting more data. Indeed, the very extensive data collection undertaken by 20 investigative journalists of the *Washington Post* over two years, and made publicly available, leaves us with an impression of partiality and imprecision mainly for this reason (Leander 2014). Instead, methodological imagination and flexibility is needed to reflect on how more persuasive data could be collected.

Analogously, also for qualitative data, the great divides run through the self-understandings, actions and stories of the researched and fashion the research process. The divides therefore take on a practical importance that often goes against the grain of the expectations of both researchers and the researched, who start out with specific assumptions about where the NAPIS belong in the divide. For example, Joakim Berndtsson (see Chapter 6) recounts how Swedishness imposed itself as a 'subtheme', quite contrary to his expectations. Elsewhere Joakim Berndtsson and Maria Stern (2011) discuss the many unexpected paradoxes, contradictions and tensions generated by the categorical divides. Methodological flexibility and openness are a precondition for allowing these kind of unexpected themes, which break the great dividing lines, to impose themselves, and indeed for contemplating how better to collect data of this kind; perhaps making unlikely methodological combinations is of essence. As Andreas Kruck (see Chapter 8) argues, although such flexibility and willingness to make unexpected combinations may 'be a cause for concern for the purist methodologist ... from the perspective of problem-oriented research ... [it] appears to be a fruitful strategy.'

Another reason the process of analysing and interpreting data pertaining to NAPIS is particularly challenging is the omnipresent and methodologically central challenge of handling the consequences of observing from a specific location: what you see, understand and say depends on where you stand. Indeed, observation necessarily and always takes place from a specific location and with the help of specific observational instruments and devices that will shape the vision of the observed. The object of observation is constructed by the observer (Bourdieu *et al.* 1991). The way this inevitable challenge is articulated and handled varies. In the hard sciences or in quantitative analysis, this makes it essential to be aware of the significance of the kind of observations involved in constructing 'facts' (Barad 2007; Latour 2010 among many). In fieldwork and interviews it becomes crucial to be aware of the ways in which the researchers and the researched interact. As Melanie Coni-Zimmer and Klaus Dieter Wolf (see Chapter 14), Tessa Diphoorn (see Chapter 12) and Anja Mihr (see Chapter 5) discuss in their contributions, this includes the place of the emotions of the researcher, but also reflections on how the presence of a researcher as an embodied personality transforms the object of observation (also Bourdieu 1999: Chapter 6; Flam and Kleres 2015). Being an unnoticed 'fly on the wall' simply is not an option.

The fact is that this challenge of handling the point of observation is particularly daunting for NAPIS research, for the simple reason that the great divides shape not only the observed but also the observer. They fashion the way researchers understand what they are observing. Concretely, whether observing quantitative data, reading texts, watching images or listening to statements, observers will hear, see, feel, understand and analyse in specific categories. The experience of going through recorded material to find that something very clearly said was overheard because it did not fit the categories of listening; being surprised at discovering an unfamiliar way of reading a text; suddenly finding that ways of doing things make sense; discovering a new way of seeing an image; or realizing that another way of organizing data, or indeed collecting it, helps show a link, all point to the importance of the unexpected. It is part of the excitement and reward involved in doing research. The specificity of NAPIS research is that, precisely because the state is an epistemological problem ‘that makes itself felt in innermost thoughts’ (Bourdieu 1994: 100), this kind of surprise has to be consciously cultivated, otherwise NAPIS will be overlooked and ignored. As Anja Mihr (see Chapter 5) puts it, the main challenge is to ‘control subjectivity’ enough to closely observe and carefully listen, so as not to miss the NAPIS, or crucial aspects of these, even when they are strikingly visible or really loud (Johns 1994; 2013, introduction). To discover the marginalization of human rights violations in UK narratives about Libyan rebels requires conscious control, offsetting the dominance of a public narrative that focuses on their heroism (see Chapter 3 by Alexander Spencer).

Furthermore, method also concerns writing: how does one communicate research to others so that they will understand what is being said and find the arguments persuasive and justified? Again, general issues that weigh on all researchers are particularly acute for NAPIS research. All researchers face choices about how much space in their text to accord their objects of research and how much to cover them up and insert them into theoretical issues; how much and what kind of violence of writing to exert. At one end of the spectrum are those who argue that poetry, art, stories or music best convey an understanding of the observed; at the other, those who include the observed as tightly constructed variables in an equation. Most, including the contributors to this volume, fall somewhere in-between, combining different forms of writing. The choices are clearly ‘personal’, although in a manner that ties into disciplinary contexts, requirements for publication, and national/institutional structures and preferences.

Communicating NAPIS research is complicated by at least two things that are specific to it. The first revolves around the ethics of research and writing about NAPIS. It is increasingly a requirement for all kinds of research that ethical guidelines and codes be followed. This is justifiable and often comforting for researchers heading out to do research (see Chapter 12 by Tessa Diphoorn and Chapter 13 by Sabrina Karim). However, if ethical guidelines and codes replace reflection on what ethical research indeed is, the good intentions can become problematic for all kinds of reasons (Schrag 2010). The lack of understanding and engagement with NAPIS makes this scenario discouragingly common in

research in the area. Requirements that might seem unproblematic in some contexts, and are standard for many journals and funding agencies, including, for example, informed consent forms, commitments to keep full records of interviews or to allow reviewing of quotes and material, can easily fetter research communication as they are de facto impossible to keep. Similarly, the sensitivity of (post-)conflict contexts places particular ethical strains on research work. These are ill captured by guidelines developed for other contexts and problems, such as those the US Institutional Review Board developed with medical research on human subjects in mind (see Chapter 13 by Sabrina Karim). A more reflected, open and imaginative approach to research ethics that can relate to the specific context, is therefore core to communication.

In addition to this, research communication – and particularly the currently dominant journal article communication – requires speaking to existing writing and literatures. This makes good sense and is intended to avoid eternal reinventions of the wheel. However, in view of the centrality of the great divides in the construction of knowledge, this entails a feat for NAPIS research. One way this is articulated is in the common concern of NAPIS researchers with bias, including their own ‘judgementalism’ (see Chapter 12 by Tessa Diphoorn), but also that of the literature (see Chapter 14 by Melanie Coni-Zimmer and Klaus Dieter Wolf). The requirement of speaking to the literature tends to impose and perpetuate this sense of bias and the difficulty of surmounting it. Hence, although NAPIS have been amply studied, research communication often does more to reinforce than to resolve the difficulty of studying them (Leander 2011).

This leads directly to the final point I wish to make, namely that the possibly most important reason for adopting a methodological perspective on NAPIS that works as an ‘encyclopedia in the making’ rather than a cookbook, is to open a route to improve the questions asked about NAPIS. The logic of the requirement to speak to the literature is not only stultifying because it imposes unhelpful terms of communication. It is sometimes misinterpreted to imply that questions, mechanisms and processes that have not already been studied cannot be studied. Unless an academic authority has named them, they cannot possibly matter. This is particularly troublesome for NAPIS research. The implication is that even when the observed is granted the space to speak back to the researcher and to be actually heard, researchers will often end up ignoring what they hear and see because it does not fit the methodological strictures. And if they don’t do it themselves, reviewers and journal editors will usually push them in that direction. The frustration vented by researchers, both junior and senior, who find the most important and interesting aspect of their research ending up as a footnote, marginalized or even invisible, addresses the point.

One way of dealing with this difficulty is to refuse communication on these terms. Susan Strange opposed what she termed the ‘Procrustean bed’ of theory, and introduced her PhD classes by explaining that the structure of research was to first say what other had said and then proceed (Chapter 2) to explain why they had it all wrong, and she devoted one of her last writings to the stifling effects of theory and method (Strange 1995, 1996, 1998 respectively). Fortunately, as this

volume testifies, there are alternatives to this decreasingly tenable refusal to engage with method. One is to find authorities who have somehow engaged with the issues, for example by shifting theoretical and disciplinary terrain. Since academic authorities have worked on most issues, this is indeed a very realistic possibility for circumventing the problem and being allowed to reformulate the problem, and no doubt a reason why NAPIS research tends to be inherently interdisciplinary (Abrahamsen and Leander 2016). Another alternative is to learn from the anthropologists and ethnologists, who have turned it into a virtue to give space to the observed, and even to reverse the perspective by relying on their categories to redefine the terms (Viveiros de Castro 1992; Strathern 1996; Riles 2002). Perhaps it is not surprising that, in this volume, the contributors who most explicitly reflect on how their engagement with their field work led them to redefine their research focus and research questions are those most clearly engaged in ethnographically inspired work (see Chapter 12 by Tessa Diphoorn and Chapter 6 by Joakim Berndtsson). Clearly, this willingness to allow the observed to redefine the questions could also be made far more general and applicable to other research traditions (Leander 2015). This would indeed be an important methodological perspective for research about NAPIS.

Conclusion

To sum up, this volume demonstrates why it is important for research focused on Non-State Actors and Processes in International Security to leave behind the cookbook methods, with recipes for secure knowledge production. As this conclusion has argued – and the contributions to this volume demonstrate – methods need to be adjusted to the issues studied and, if well used, they can play an important role in redefining these issue by directing attention to overlooked mechanisms, processes, issues actors, logics, rationalities, or technologies. This requires an approach that looks at methods as an ‘encyclopaedia in the making’, where researchers are not merely ‘applying methods’ but adjusting them and rewriting them as they go along; an approach that leaves space for combinations, imagination and improvisation but also for learning from the mistakes of others.

Much of this concluding chapter has been devoted to highlighting the reasons why this approach is particularly important for NAPIS research, which is what this volume is devoted to. NAPIS research crosses what I have termed the great divides: inside/outside, public/private, civil/military, state/market, created by the historically complex and contextually varied emergence of modern social political thought that has stabilized the state’s monopoly on symbolic violence to which these categorical divides are pivotal. NAPIS research is therefore particularly subjected to the state as an epistemological problem. Researching NAPIS demands not only that we resist the way these divides hamper NAPIS research by shaping data collection, data interpretation, writing, and the kinds of questions that can be asked. It demands that we incorporate them into the studies so as to make their performative effects visible. Such incorporation involves reconquering ‘a right to look’ at and problematize the consequences of the dominant

point of view and the principles of vision and division in it (Mirzoeff 2011). One of the main promises of NAPIS research is that it is often – perhaps unwittingly and unwillingly – pushed to do this and to discuss how it was done. Herein lies, I think, the methodological perspective of NAPIS research, which is of relevance not only for the growing NAPIS research community but for scholarship, research and the critical political and social theory well beyond it.

Notes

- 1 I add processes since for many approaches (including practice, system, structuralist, or discourse analytical approaches) the processes are at least as important as the actors.
- 2 I prefer this reference to a Wikipedia-like encyclopaedia to existing alternatives, such as for example Latour's (2005: 17) travel guide image, precisely because it emphasizes researchers' involvement and the constant change.

References

- Abbott, Andrew. 2004. *Methods of Discovery: Heuristics for the Social Sciences*. New York and London: W.W. Norton and Company.
- Abrahamsen, Rita, and Anna Leander. 2016. 'Introduction', pp. 1–9 in *Routledge Handbook in Private Security Studies*, R. Abrahamsen and A. Leander, eds. London: Routledge.
- Aradau, Claudia, and Jef Huysmans. 2014. 'Critical Methods in International Relations: The Politics of Techniques, Devices and Acts'. *European Journal of International Relations* 20(3): 596–619.
- Barad, Karen. 2007. *Meeting the Universe Half-Way: Quantum Physics and the Entanglement of Matter and Meaning*. Durham and London: Duke University Press.
- Berndtsson, Joakim, and Maria Stern. 2011. 'Private Security and the Public-Private Divide. Contested Lines of Distinction and Modes of Governance in the Stockholm-Arlanda Security Assemblage'. *International Political Sociology* 4(5): 408–425.
- Boltanski, Luc. 2014. *Mysteries and Conspiracies: Detective Stories, Spy Novels and the Making of Modern Societies*. Oxford: Polity Press.
- Bourdieu, Pierre. 1994. *Raisons pratiques. Sur la théorie de l'action*. Paris: Seuil.
- Bourdieu, Pierre, ed. 1999. *The Weight of the World: Social Suffering in Contemporary Society*. Cambridge: Polity.
- Bourdieu, Pierre. 2003. *Meditations Pascaliennes*. Paris: Editions du Seuil.
- Bourdieu, Pierre. 2012. *Sur l'Etat: Cours au Collège de France 1989–1992*. Paris: Editions du Seuil.
- Bourdieu, Pierre, Jean-Claude Chamboredon, Jean-Claude Passeron, and Beate Kraus. 1991. *The Craft of Sociology: Epistemological Preliminaries*. Berlin and New York: De Gruyter (original edition 1968).
- Czarniawska, Barbara. 2016. 'The Fate of Counter-Narratives: In Fiction and in Actual Organizations', pp. 195–207 in *Counter-Narratives and Organization*, T. Kuhn, M. Wolff Lundholt, and S. Frandsen, eds. New York: Routledge.
- Elshtain, Jean Bethke. 1981. *Public Man, Private Woman: Women in Social and Political Thought*. Oxford: Martin Robertson (second edition with Princeton University Press, 1993).
- Feyerabend, Paul. 2010. *Against Method*. London: Verso.

- Flam, Helena, and Jochen Kleres, eds. 2015. *Methods of Exploring Emotions*. New York: Routledge.
- Gadamer, Hans-Georg. 1990 [1960]. *Wahrheit und Methode. Grundzüge einer philosophischen Hermeneutik*. Tübingen: J.C.B. Mohr.
- Graz, Jean-Christophe. 2008. 'Gare aux hybrides: Mythes et réalités de la gouvernance de la mondialisation'. *Etudes internationales* XXXIV(3).
- Hurt, Shelley, and Ronnie Lipschutz, eds. 2016. *Hybrid Rule and State Formation: Public-Private Power in the Twenty-First Century*. London and New York: Routledge.
- Johns, Fleur. 1994. 'The Invisibility of the Transnational Corporation: An Analysis of International Law and Legal Theory'. *Melbourne University Law Review* 19: 893–923.
- Johns, Fleur. 2013. *Non-Legality in International Law: Unruly Law*. Cambridge, UK: Cambridge University Press.
- Latour, Bruno. 2005. *Re-assembling the Social. An Introduction to Actor Network Theory*. Oxford: Oxford University Press.
- Latour, Bruno. 2010. *On the Modern Cult of the Factish Gods*. Durham, NC: Duke University Press.
- Law, John. 2004. *After Method, Mess in Social Science Research*. London: Routledge.
- Leander, Anna. 2011. 'Risk and the Fabrication of Apolitical, Unaccountable Military Markets: The Case of the CIA "Killing Program"'. *Review of International Studies* 37(5): 2253–2268.
- Leander, Anna. 2014. 'Understanding U.S. National Intelligence: Analyzing Practices to Capture the Chimera', pp. 197–221 in *The Return of the Public in Global Governance*, J. Best and A. Gheciu, eds. Cambridge: Cambridge University Press.
- Leander, Anna. 2015. 'Strong Objectivity in Security Studies: Ethnographic Contributions to Method Development'. *International Studies Perspectives*. doi: <http://dx.doi.org/10.1093/isp/ekv021>.
- Leander, Anna. 2015–2016. 'Mercados transgresores de seguridad: una mercancía en disputa y sus prácticas de mercado | Transgressive Security Markets: A Contested Commodity and its Market Practices'. *Relaciones Internacionales* (30): 117–137.
- Mirowski, Philip. 2011. *Science-Mart: Privatizing American Science*. Harvard: Harvard University Press.
- Mirzoeff, Nicholas. 2011. *The Right to Look : A Counterhistory of Visuality*. Durham, NC: Duke University Press.
- Nowtny, Helga. 2016. *The Cunning of Uncertainty*. Cambridge, UK: Polity Press.
- Przeworski, Adam, and Frank Salomon. 1995. *On the Art of Writing Proposals: Some Candid Suggestions for Applicants to Social Science Research Council Competitions*. New York: The Social Science Research Council.
- Riles, Annelise. 2002. 'Rights Inside Out: The Case of the Women's Human Rights Campaign'. *Leiden Journal of International Law* 15(2): 285–305.
- Ruppert, Evelyn, John Law, and Mike Savage. 2013. 'Reassembling Social Science Methods: The Challenge of Digital Devices'. *Theory, Culture & Society* 30(4): 22–46.
- Salter, Mark, and Can Mutlu, eds. 2012. *Research Methods in Critical Security Studies: An Introduction*. London: Routledge.
- Savage, Mike. 2013. 'The "Social Life of Methods": A Critical Introduction'. *Theory, Culture & Society* 30(4): 3–21.
- Scholte, Jan Aart. 2005. *Globalization: A Critical Introduction*. London and New York: Palgrave Macmillan.
- Schrag, Zachary M. 2010. *Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965–2009*. Baltimore, MD: John Hopkins University Press.

- Shapiro, Michael J. 2012. *Studies in Trans-Disciplinary Method: After the Aesthetic Turn*. London: Routledge.
- Steinmetz, Georg, ed. 2005. *The Politics of Method in the Human Sciences: Positivism and its Epistemological Others*. Durham and London: Duke University Press.
- Stengers, Isabelle. 1995. *L'invention Des Sciences Modernes*. Paris: Flammarion (Original edition 1993).
- Strange, Susan. 1995. 'The Defective State'. *Daedalus* 24(2): 55–74.
- Strange, Susan. 1996. *The Retreat of the State: The Diffusion of Power in the World Economy*. Cambridge, UK: Cambridge University Press.
- Strange, Susan. 1998. 'What Theory? The Theory in Mad Money'. *Centre for the Study of Globalization and Regionalism*. Working paper (18), University of Warwick, UK.
- Strathern, Marilyn. 1996. 'Cutting the Network'. *The Journal of the Royal Anthropological Institute* 2(3): 517–535.
- Tronto, Joan. 2013. *Caring Democracy. Markets, Equality and Justice*. New York: New York University Press.
- Ulmer, Jasmine B., and Mirka Koro-Ljungberg. 2015. 'Writing Visually Through (Methodological) Events and Cartography'. *Qualitative Inquiry* 21(2): 138–152.
- Viveiros de Castro, Eduardo. 1992. *From the Enemy's Point of View. Humanity and Divinity in an Amazonian Society*. Chicago: Chicago University Press.
- Walker, R.B.J. 1988. *One World, Many Worlds: Struggles for a Just World Peace*. New York: Longman.
- Wolin, Sheldon S. 1969. 'Political Theory as Vocation'. *The American Political Science Review* 63(4): 1062–1082.

Index

Page numbers in *italics* denote tables, those in **bold** denote figures.

- abstract concepts, quantifying 152
- Academi Elite *see* social media study
- academic consensus approach, to study of terrorism 15, 20–2
- access, for research 4, 5, 180, 210–11, 213, 214, 226
- access to data: security issues 102; *see also* private security study
- accountability, private military and security companies (PMSCs) 114
- action, and characterization 35
- active participation 174
- actors, state and non-state 4
- administrative units, geospatial analysis 147–9
- ambiguity, empirical results 152
- Amnesty International 41, 42
- analogies 33
- armed conflicts, range of participants 14–15
- armed response sector, South Africa 172–3
- authenticity, field experiments 191–2
- authority 235

- Bennett, A. 52, 116
- Bergman, A. 93n1
- Bernard, H.R. 65, 67
- between method triangulation 211
- bias: acknowledging 103; as concern 240; data sources 150; in quantitative observational studies 188–90; selection bias/effect 224–5; social desirability bias 196–7; survey experiments 196–7
- Black, D. 17
- Blatter, J. 112, 117
- Blitzkrieg fieldwork 179
- Blume, T. 112, 117
- Bogner, A. 78
- book: aim 1; approach taken 1–2, 9, 10, 222; chapter overviews 4, 5, 6–7; methods covered 2; questions asked 2; scope 6, 8–9; structure 9–10; summary and conclusions 241–2; *see also* NAPIS research
- Boole, G. 125
- Boolean algebra, qualitative comparative analysis (QCA) 125–6
- Boolean notation 125–6
- boundaries: contextual 102–3; NAPIS research 235–6
- Bourdieu, P. 236, 239
- Braun, V. 86, 91, 100
- Bryman, A. 52, 85

- calibration, fuzzy sets 125
- Cameron, D. 43
- Carmola, K. 83, 86, 91
- case selection 110, 209–11, 226
- case study analysis: reflexivity 226–8; *see also* corporate engagement
- casualty-sensitivity 114–15
- causal asymmetry 127
- causal claims assessment: combined case study techniques 161; context and overview 159–60; data analysis method 164–6; data collection 163–4; defining causality 165; geospatial analysis 162; knowledge generation 166; method selection 161–2; process tracing 164–5; qualitative comparative analysis (QCA) 162, 165; questions asked 159; summary and conclusions 166–7; units of analysis 164; *see also* combined case study techniques; geospatial analysis; qualitative comparative analysis (QCA)

- causal emplotment 35
- causal process tracing 116, 117–18, 120
- causality: defining 165; perspectives on 159–60; time sequences 160
- causality loops 189
- characterization 34–5; Libya study 37–40
- chimera analogy 236
- Clarke, V. 86, 91, 100
- classic research design, dependent variable 24
- codebook, corporate engagement study 209
- cognitive perspective, narrative analysis 33
- Collier, D. 116
- colloquialisms 101
- combined case study techniques: aim of study 107–8, 109; case selection 110; challenges 119–20; comparative interpretation 118; congruence analysis 120; context and overview 107–9; cost-efficiency hypothesis 114; covariational analysis 111–16, 119–20; data sources 110–11; dependent variables 112–13; descriptive findings 113; explanatory and theoretical models 108; focus of study 208–9; functional complexity hypothesis 113–14; independent variables 113; measurement of security privatization 110–11; operationalization 110–11; political model 114–15; process tracing 120; research design 109–19; research question 108; summary and conclusions 119–20; triangulation 108; *see also* causal claims assessment
- communication 239–41
- comparative analysis 227
- comparative interpretation 118
- complex causation 126–7
- conceptualization: academic consensus approach 20–2; conceptual frameworks 15–24; constructivist approach 16–17; context and overview 14–15; criminological approach 17–18; enumerative definitions of terrorism 20–1; gradual modelling approach 21–4, 23; legal approach 18–20; potential research questions 24; purpose of defining terrorism 19; research scenarios 24–5; summary and conclusions 25
- congruence analysis 116–18, 120
- conjunctural causation 123, 127
- consistency, measures of 127–8
- constructivist approach 16–17; to study of terrorism 15–16; understanding of politics 33; use of legal definitions 19; *see also* social constructivism
- content analysis 25, 97, 100–1, 212; *see also* social media study
- contextual knowledge 101
- cookbook approach, vs. interactive encyclopaedia approach 233–4
- coping strategies 8
- corporate engagement: access 209–10, 213, 226; aims of study 205; approach taken 205; case selection 209–11; codebook 209; conceptual issues 226–7; context and overview 204–5; corporate governance contributions 205; corporate outcomes 207–8; corporate output 206–7; data collection 211; dimensions of effectiveness 206; document content analysis 212; effectiveness 206–9; ethics 214–15, 227–8; expert interviews 212–13; field research 213–16; impact 208, 227–8; methods 211–16; reflexivity 226–8; research question 204; results 216; summary and conclusions 216–17; triangulation 205, 227
- corporate social responsibility (CSR) 204, 205
- cost-efficiency hypothesis 114
- covariational analysis 111–16, 119–20
- coverage, measures of 127–8
- credibility 235
- crime, compared with terrorism 17–18
- criminal law, and terrorism 18
- criminological approach, to study of terrorism 15, 17–18
- crisp sets 124–5
- critical feminism 223
- critical theory, on causality 160
- crowdsourcing, data sources 150
- cultural competence 101, 103–4
- cultural memory 33
- cultural perspective, narrative analysis 33
- Darfur: data sources 149–50; possible causes of conflict 143; remote sensing data 153–4; state use of militias 143; theoretical argument 151; *see also* geospatial analysis
- data: problems of lack of 152–3; remote sensing data 153–4
- data collection: challenges of 98–9; experimental methods 199–200; NAPIS research 237–8
- data sources: bias 150; combining 99; crowdsourcing 150; geospatial analysis 150; Libya study 37, 39; reliability 4, 5, 85; social media study 51

- data triangulation 211
- Deitelhoff, N. 205
- Denzin, N. 97–8, 211
- dependent variable: classic research design
 - 24; construction of 25
- DeWalt, B. 173, 174, 178, 179, 180
- DeWalt, K. 173, 174, 178, 179, 180
- Diphorn, T. 182
- direct speech, and characterization 35
- do no harm principle 197, 198
- documents: redaction 87–8; reliability 85
- ecological inference fallacy 144–5
- economic freedom 115
- effectiveness: concept of 205, 206;
 - corporate engagement 206–9;
 - dimensions of 206
- emotion: Libya study 37; when researching violence 176
- empirical results, ambiguity 152
- employment: causal 35; Libya study 38–9, 40
- endogeneity 189, 224
- endorsement experiments 196
- Enloe, C. 227
- enumeration teams 199–200
- enumerative definitions, of terrorism 20–1
- environmental and livelihoods
 - vulnerability mapping (ELVM) 149–51, 154
- equifinality 123, 127
- ethical guidelines 89
- ethics 197–200, 214–15, 227, 239–40
- Evans-Pritchard, E. 175
- experimental methods: challenges 199;
 - comparison of methods 191–200;
 - context and overview 187–8; data collection 199–200; endorsement experiments 196; enumeration teams 199–200; ethics 197–200; field experiments 191–4; informational survey experiments 195; lab-in-the-field 194–5; list experiments 196; vs. observational approaches 188–91; pre-analysis plans 197–8; random assignment 190, 198–9; reflexivity 223–5; reporting back 200; safety 197, 199; summary and conclusions 200–1; survey experiments 195–7
- experts, interviewing 69–70, 74
- explanatory conditions 131–3
- external validity, field experiments 193
- feminist standpoint theory 230n1
- field experiments 191–4; authenticity
 - 191–2; context 192; disadvantages and trade-offs 193–4; external validity 193;
 - lab-in-the-field 194–5; outcomes 192–3;
 - participants 192; reflexivity 223–5;
 - resources 193–4
- field notes, participant observation 178–9
- field research: corporate engagement
 - 213–16; ethics 214–15; *see also* experimental methods; participant observation
- flexibility, methodological 238
- Fludernik, M. 33, 35
- force, legitimacy of 15
- Forsey, M. 174, 177
- framing 50
- France *see* combined case study techniques
- Freilich, J.D. 17
- functional complexity hypothesis 113–14
- funding 235
- fuzzy sets 123, 124–5
- Gadamer, H.-G. 234–5
- Gaddafi, Col. M. *see* Libya study
- Gallie, W.B. 16, 20
- gender: participant observation 228–9; and reporting of sexual violence 189–90; *see also* experimental methods; sexual and gender based violence (SGBV)
- geographical disaggregation 153
- George, A. 116
- geospatial analysis: administrative units
 - 147–9; context and overview 143–4;
 - data sources 149–50; designing 145–55;
 - environmental and livelihoods vulnerability mapping (ELVM) 149–51, 154; geographical disaggregation 153;
 - grid cells 147–9; hypothesis formulation 145–6; interpreting results 154–5;
 - measuring explanatory variables 151–4;
 - measuring outcome variables 149–51;
 - method 143–4; multi-method design 154–5; operationalizing outcome variable 151; problems of lack of data 152–3; questions asked 144;
 - recommendations 155–6; remote sensing data 153–4; software 143;
 - spatial disaggregation 145; subnational geospatial analyses 144–5, 146;
 - summary and conclusions 155–6;
 - theoretical argument 151; units of analysis 146, 147–9, 147; *see also* causal claims assessment; Darfur
- Gerber, A. 191

- Germany: human rights awareness 69; *see also* combined case study techniques
- Gerring, J. 110, 112
- Global Terrorism Database (GTD) 21
- Goertz, G. 134
- Goldstein, D. 171
- governance: market-based 115; role of non-state actors 204–5; types of contribution to 205
- gradual modelling approach **23**; potential research questions 24; research scenario 24–5; to study of terrorism 15, 21–4
- Green, D. 191
- grid cells, geospatial analysis 147–9
- grounded theory 229
- Hague, W. 43
- Hammond, P. 43
- hand of God 198
- hanging out 175–7
- Hayes, J. 16
- Herrnstein Smith, B. 32
- historical narratives, private military and security companies (PMSCs) 117–18
- how of research *see* method; methods
- human rights, awareness survey 69
- Human Rights Watch 41
- hybridity, private security companies (PSCs) 83, 91
- hypothesis formulation 133–4
- ideal types 22
- ideational approach 108–9, 115
- identity construction, through narrative 33–4
- indifference, terrorism 17–18, 22, 24–5
- Indonesia, data sources 150
- informational survey experiments 195
- Ingold, T. 174
- innovation, opportunities for 5–6
- Institutional Review Boards (IRBs) 197–8
- interdisciplinary research 66
- intermediate solutions 135
- interpretive analysis: boundaries 102–3; content analysis 100–1; contextual knowledge 101; cultural competence 101; limits of knowledge 102; narrative analysis 99–104; participant attributes 102–3; road maps 100–1; scripting 100–1; sentiment analysis 100, 101, 102; thematic analysis 100
- intersubjectivity 223
- interview guides 66
- interviewees: choice of 71–5; expectations/motivations 215; motivation for participation 73–4, 78
- interviews: corporate engagement study 212–13; data created 66; with experts 69–70, 74; interviewer knowledge 77; language issues 71, 73, 74–5; neutrality 77; non-verbal cues 73; open interviews 177; particularities of 98; preparation for 78; private security study 88–90; probing 77–8; sensitivity 72; subjectivity 76; types of 65–6; use of 65
- INUS and SUIN causes 127, 127
- investigator triangulation 211
- Iraq War study: analysis of necessary conditions 134–5; calibrating explanatory conditions 131; context and overview 130; fuzzy set calibration 133; fuzzy set, military participation 131–3, 132; hypothesis formulation 133–4; interpretation of results 139; military participation and solution term **138**; results visualization 138–9; selecting explanatory conditions 131–3; solution terms 135, 137–8, 138; specifying dependent variable 130; truth table analysis 135, 136; variance in military contributions 130; *see also* qualitative comparative analysis (QCA)
- Jackson, M. 182
- judgementalism 176
- Kinsey, C. 89
- knowledge: limits of 102; and practice 1
- Krasner, S.D. 206
- lab-in-the-field experiments 194–5
- labelling 35
- LaFree, G. 17
- language issues: case selection 210; colloquialisms 101; interviewing 71, 73, 74–5; participant observation 180
- Leander, A. 86
- Lee, L. 101, 102
- legal approach, to study of terrorism 15, 18–20
- legitimacy: of force 15; of terror 20, 21–2
- Libya study: characterization 37–8, 39–40; data sources 37, 39; emotion 37; emplotment 38–9, 40; media narratives 37–9; narratives of political elite 39–40; overview 36; setting 37; settings 39; war crimes and human rights violations 41–3; *see also* narrative analysis
- limited diversity 129–30, 131, 135
- Lincoln, Y.S. 97–8

- list experiments 196
- listening: participant observation 177, 229; semi-structured interviews 77
- Liu, B. 51
- local knowledge, using 198
- location effects 238
- logical minimization, of truth tables 128–9
- longitudinal research, participant observation 179–80
- luck 235
- machine learning 54–5
- Madden, R. 181
- Malinowski, B. 173, 175
- mapping the scene 178
- measures of consistency, qualitative comparative analysis (QCA) 127–8
- measures of coverage 127–8
- Mello, P.A. 134
- metaphors 33
- method: categorical division 237; choice of 1, 96; engaging with 240–1; flexibility 238; redefinitions 234–5; writing 239–41; *see also* NAPIS research
- method as closure 234–5
- methodological nationalism 236
- methodological openness 237
- methodological vocabularies 235
- methodology, politics of 4
- methods: answering and asking questions 237–41; choice of 5–6, 7–8; and credibility 235; importance of 234–7; selectivity 6; *see also* NAPIS research
- methods triangulation 211
- mixed methods 6–7, 97; supplementing data 68; *see also* social media study
- modifiable areal unit problem (MAUP) 147
- Monroe, K. 100, 101
- Moon, D. 99–100
- moral assessment 20, 21–2
- mosaics of data 178
- multi-method designs, geospatial analysis 154–5
- Murray, S. 178
- mutual constitution 160
- myths 33
- naming 35
- NAPIS research 233–4; answering and asking questions 237–41; categories and boundaries 235–6; context 235–6; context and overview 233; cookbook approach vs. interactive encyclopaedia approach 233–4; data collection 237–8; engaging with method 240–1; ethics 239–40; importance of methods 234–7; improving questions 240; location effects 238; observer effect 238–9; summary and conclusions 241–2; writing 239–41
- narrative analysis 96, 100; causal emplotment 35; characterization 34–5; cognitive perspective 33; constituent elements 34–5; context and overview 31; as critical method 41–3; cultural perspective 33; design and conduct 35–6; relevance 33; settings 34; software 36; summary and conclusions 43; temporality 35; usefulness 31; *see also* Libya study
- narratives: concepts of 32; historical 117–18; identity construction 33–4; importance of 31; relevance 32–3
- narratology 34
- natural language processing 52
- necessary conditions 134–5
- neoliberalism 115–16
- nested analyses 156
- Neumann, I. 98, 100, 101, 103–4
- neutrality 77, 215
- Nigeria, data sources 150
- Nilan, P. 182
- Nisour Square 51, 59
- non-outcomes 139
- non-random refusal 196–7
- non-state actors: data sources 4; governance role 204–5; political agendas 5; purpose of interviews 67–70; research foci 3; as research objects 2–4; and states 4
- Non-State Actors and Processes in International Security (NAPIS) research *see* NAPIS research
- non-verbal cues 73
- Normalized Difference Vegetation Index (NDVI) 153–4
- notes, re-reading 179
- objectivity 223
- observation 177–8; *see also* participant observation
- observer effect 223, 238–9
- open interviews, participant observation 177
- open mindedness 89, 103
- operationalization, combined case study techniques 110–11

- opinion mining *see* sentiment analysis
- outcome effectiveness 205
- outcomes, defining 206, 226–7
- output effectiveness 205
- outputs, defining 206

- Pakistan, data sources 150
- Palme, O. 16
- Pang, B. 101, 102
- parliamentary control, private military and security companies (PMSCs) 114
- participant attributes, effects of 102–3
- participant observation: access 180; advantages 182; applications 174–5; armed response sector project 172–3; context and overview 171–2; as continuum 174; defining 173–5; elements 172, 175–83, 228; emotional experiences 176; the field 175; field notes 178–9; gender 228–9; language issues 180; listening 177, 229; longitudinal research 179–80; mapping the scene 178; methods 173; modes of participation 174; observation 177–8; open interviews 177; participation and hanging out 175–7; public-private security assemblages project 173; rapport 180–1; reflexivity 181–2, 228–9; research projects 172–3; summary and conclusions 183–4; trust 180–1; what people say and do 182–3
- participatory observation 65–6, 73
- passive participation 174
- Patterson, M. 33, 100, 101
- Peace Research Institute Frankfurt 204
- Phelan, J. 32
- policy research *see* corporate engagement
- political psychology 187
- political violence: in classic research design 24; *see also* terrorism
- politics, of methodology 4
- politics of, constructivist approach 33
- post-positivism 223
- power-based approach 108, 114–15
- pre-analysis plans 197–8
- preconceptions 103–4
- private military and security companies (PMSCs): accountability 114; congruence analysis 116–18; expansion of role 50; historical narratives 117–18; ideational approach 108–9, 115; and neoliberalism 115–16; parliamentary control 114; power-based approach 108, 114; process tracing 116, 117–18; research into 50; and transparency 114; use of social media 50–1; *see also* combined case study techniques; private security study; social media study
- private security companies (PSCs): as hybrids 83, 91; use of 81
- private security study: access to data 84–5, 93n2; aim 82, 84; consent 89; context and overview 81–2; data analysis 90–2; data collection 85–6, 87–90; data reliability 85; design 82–7; discussion and conclusions 92–3; ethics 89; further research 93; images presented 84; interviews 88–90; necessary scope 83; possible alternative methods 92; prior knowledge 82; questions addressed 81–2; redaction of documents 87–8; settings 84; single case study 83–4; thematic analysis 82, 86, 90–2; themes and constructs 91–2; theoretical bases 82–3; triangulation 93; *see also* private military and security companies (PMSCs)
- probing 77–8
- process tracing 116, 117, 120
- professional image, private security study 84
- Public Goods Game 194–5
- public image, private security study 84
- public-private security assemblages project 173

- qualitative comparative analysis (QCA) 111–12; advantages 124; applications 123–4; Boolean algebra 125–6; causal asymmetry 127; complex causation 126–7; conjunctural causation 123, 127; context and overview 123–4; crisp-set 124–5; equifinality 123, 127; fuzzy-set 124–5; fuzzy set calibration 125; INUS and SUIN causes 127, 127; limited diversity 129–30; measures of consistency 127–8; measures of coverage 127–8; principles and terminology 124–30; set-theoretic methods 126; software 134; statistical estimation techniques 125; summary and conclusions 140; truth tables 128–9, 129; *see also* causal claims assessment; Iraq War study
- qualitative content analysis, social media study 58–60
- quantitative observational studies 188–9
- questions, answering and asking 237–41

- Quine-McCluskey algorithm 128–9
- Ragin, C. 123, 125, 128
- random assignment 190, 198–9
- randomized control trials (RCTs) 187
- rapport, participant observation 180–1
- rational choice theory 187
- readability 112
- redaction, of documents 87–8
- reflexivity: bias 224–5; case study analysis 226–8; experimental methods 223–5; objectivity/subjectivity 223; overview 222; participant observation 181–2, 228–9; research as social 222–3; summary and conclusions 229–30
- regime theory 205
- regimes, defining 206
- reluctant participation 174
- remote sensing data 153
- Renwick Monroe, K. 33
- reporting back 200
- research: challenges 3–6; opportunities for innovation 5; as social 222–3
- research access 180
- research scenario, gradual modelling approach 24–5
- results, reporting back 200
- Richardson, L. 22
- Rigako, G. 175
- road maps 100–1
- rule of law, violations 18
- Sack, F. 19
- safety, experimental methods 197, 199
- scaling 147
- Schmid, A.P. 20, 21
- science, as contested 235
- scripting 100–1
- security, research challenges 4–5
- Security for Sale project 81; *see also* private security study
- security governance 205
- security issues: access to data 102; *see also* private security study
- security privatization *see* private military and security companies (PMSCs)
- security sector reform, dearth of literature 188
- selection bias/effect 189–90, 224–5
- semi-structured interviews 96–7; analysis phase 76; choosing interviewees 71–5; context and overview 65–7; corporate engagement study 212–13; data created 66; doing phase 76; inside perspective 92; interviewer knowledge 77; listening 77; methodological challenges 70–8; multi- and cumulative causal relationships 75–6; neutrality 77; phases 76–7; planning phase 76; preparation for 78; purpose of 67–70; recruiting interviewees 71; reflection phase 76–7; research challenges 66–7; risk of over-interpretation 70; summary and conclusions 79; types of questions 66; utility of data 75; ‘why’ question 75–8
- sense making 33
- sensitivity, when interviewing 72
- sentiment, types of 101
- sentiment analysis 51–2, 100, 101, 102; social media study 54–7
- set-theoretic methods 126
- set theory 123
- settings: Libya study 37, 39; of narratives 34; private security study 84
- sexual and gender-based violence (SGBV) 224–5
- Silke, A. 21
- single case studies 83; *see also* private security study
- situatedness 223, 228–9
- snowball sampling 213
- social constructivism: on causality 159–60; *see also* constructivist approach
- social desirability bias 196–7
- social media: research into 50; use by private military and security companies (PMSCs) 50–1
- social media study: approach taken 51; choice of subject 51; context and overview 50–2; data retrieval and preparation 52–3; data sources 51; distribution of sentences per tweet 55; method 51–2; qualitative content analysis 58–60; results of sentiment analysis 57–8; sentiment analysis 51–2, 54–7; sentiment scores as percentages of tweets 57; summary and conclusions 60–1; tweets as data 53–4
- soft triangulation 211
- software, narrative analysis 36
- solution terms, Iraq War study 135, 137–8, 138, 138
- Sommer, G. 69
- spatial disaggregation 145
- speech, and characterization 35
- Stanford NLP 55–7
- Stanford Sentiment framework 52, 55
- statistical analysis 25

- statistical correlations, limitations of 154
- statistical estimation techniques, qualitative comparative analysis (QCA) 125
- Stellmacher, J. 69
- Stern, M. 81
- stories 33
- Strange, S. 240–1
- structured interviews 65–6
- subjectivity 76, 103, 239
- survey experiments 195–7
- Sutherland, E.H. 17
- Sweden *see* private security study
- symbolic violence 236

- Tedlock, B. 182
- temporality, narrative analysis 35
- terrorism: compared with crime 17–18; conceptual frameworks 15–24; and criminal law 18; enumerative definitions of 20–1; indifference 17–18, 22, 24–5; legal definitions 18–20; legitimacy 20, 21–2; purpose of defining 19; revised academic consensus definition 20–1; use of term 16; violations of rule of law 18; *see also* conceptualization
- Terrorism and Political Violence* 17
- text: types of 97; uniqueness 98
- the field, understandings of 175
- ‘The Impact of Transitional Justice in Democratic Institution Building’: research challenges 66–7; *see also* semi-structured interviews
- thematic analysis 82, 86, 90–2, 96, 100; *see also* private security study
- theory triangulation 211
- time sequences, and causality 160
- Toft, M. 188
- Toolan, M. 32, 34
- transdisciplinary research 66
- transparency, and private military and security companies (PMSCs) 114
- triangulation 6, 211; combined case study techniques 108; corporate engagement study 205, 227; private security study 93; research benefits of 97–8; use of texts 97
- trust, participant observation 180–1
- Truth and method* (Gadamer) 234
- truth table algorithm 128–9
- truth tables 128–9, 129; analysis, Iraq War study 135, 136; logical minimization of 128–9
- Turner, M. 33
- tweets: as data 53–4; as text 97–8
- twilight policing 183, 228
- Twitter *see* social media study

- Ultimatum Game 195
- UN High Commissioner for Human Rights (UNHCR) 42
- uncertainty 235
- United Kingdom *see* combined case study techniques
- United Kingdom Terrorism Act 18–19
- United Nations Conventions on Terrorism 18
- United States *see* combined case study techniques
- units of analysis 147; geospatial analysis 146, 147–9
- universal laws 159–60

- validity 235; field experiments 193
- Van Evera, S. 116
- Van Maanen, J. 179, 182
- van Teijlingen, E. 76, 77
- veil of ignorance 102
- violence: nature of 14–15; researcher emotions 176; sexual and gender based 224–5; symbolic 236
- visibility 236
- Voluntary Principles on Security and Human Rights 205, 207, 208

- Walt, S.M. 16
- Waltz, K. 159
- Weber, M. 22
- websites, reliability 85
- Weinberg, L. 16, 20
- ‘why’ question 75–8
- within-method triangulation 211
- Wolf, K.D. 205
- Wolin S. 234–5
- Wood, E.J. 214–15
- writing 239–41

- Zadeh, L. 125
- zoning 147