Rhetorics of Radicalism *

Daniel Karell
Division of Social Science
New York University Abu Dhabi

Michael Freedman
Department of Political Science
Massachusetts Institute of Technology

June 22, 2019

For published version, see *American Sociological Review*, 84 (4), 2019.

*The authors thank Reza Hussaini, Wasim Janjua, Majid Mohammadi, Abdul Rahim, and many others who helped make data collection in Afghanistan, Iran, and Pakistan productive and safe. We also thank Ewaz Ali Bahrami, Jason Chu, Kevan Harris, Anand Gopal, Haseebullah Juma Khan, Sebastián Rojas Cabal, Brandon Stewart, Katherine Stovel, Alex Strick von Linschoten, Muhammad Usman, and Azhar Yerzhanova for their assistance and insights throughout the process of analyzing and writing. We greatly benefited from advice offered by Elisabeth Anderson, Paul DiMaggio, Noam Gidron, Steven Pfaff, Malte Reichelt, and Blaine Robbins during the later stages of the project. Finally, we are indebted to three anonymous reviewers and the editors for suggestions that were of immense help in furthering our study. Any remaining errors and omissions are our responsibility. This project was supported by Daniel Karell's research grant from New York University Abu Dhabi's Division of Social Science and the Fung Fellowship at Princeton University's Institute for International and Regional Studies. Michael Freedman received support from the Israel Institute during the project. Daniel Karell, the corresponding author, can be reached at daniel.karell@nyu.edu and PO Box 129188, Abu Dhabi, United Arab Emirates.

Abstract

What rhetorics run throughout radical discourse, and why do some gain prominence over others? The scholarship on radicalism largely portrays radical discourse as opposition to powerful ideas and enemies, but radicals often evince great interest in personal and local concerns. To shed light on how radicals use and adopt rhetoric, we analyze an original corpus of more than 23,000 pages produced by Afghan radical groups between 1979 and 2001 using a novel computational abductive approach. We first identify how radicalism not only attacks dominant ideas, actors, and institutions using a rhetoric of subversion, but also how it can use a rhetoric of reversion to urge intimate transformations in morals and behavior. Next, we find evidence that radicals' networks of support affect the rhetorical mixture they espouse, due to social ties drawing radicals into encounters with backers' social domains. Our study advances a relational understanding of radical discourse, while also showing how a combination of computational and abductive methods can help theorize and analyze discourses of contention.

Key terms: Radicalism; Textual analysis; Computational abductive analysis; Sociology of the future; Afghanistan

1 Introduction

Scholarship on radicalism depicts a heterogeneous landscape of radical messages. Radicals may, for example, champion the protection of vanishing traditions (Calhoun 2012), or express intolerance of existing social and political systems (Nasr 2005; Achilov and Sen 2017). Some radicals may justify violence against innocents (Moskalenko and McCauley 2009; Snow and Cross 2011; Rink and Sharma 2018) and glorify death (Roy 2017), whereas others may pursue nonviolent extra-institutional action (Dalgaard-Nielsen 2010; Borum 2011; Bartlett and Miller 2012; Beck 2015).

Yet, despite recognizing this diversity, the literature subsumes radical ideas and arguments, or rhetorics, under a single broad notion: radicalism entails opposition against hegemonic sociopolitical systems or powerful people, organizations, and institutions in a particular context (*e.g.* Goldstone, Gurr, and Moshiri 1991; Ferree 2003; Calhoun 2012; Bartlett and Miller 2012; Beck 2015; della Porta 2018). In an empirical parallel, the kind of radical that has lately attracted the most attention, so-called global jihadists, are nearly always portrayed as advocates of violence against external enemies (*e.g.* Benjamin and Simon 2002; Gambetta and Hertog 2016; Walter 2017b; Rink and Sharma 2018; Mitts 2019). However, recent events suggest radicalism may comprise a range of rhetorics. The Islamic State, for instance, displayed a fixation with intimate matters (Tezcür and Besaw 2017), including the notorious policing of sexual identity (Saul 2015), women’s dress (Callimachi 2016), and smoking (Callimachi 2017).

What common rhetorics might run throughout radical discourse? Because the scholarship on radicalism does not primarily focus on the nature of its discourse—unlike recent studies of nationalism and populism (Bonikowski and DiMaggio 2016; Bonikowski and Gidron 2016; Brubaker 2017)—our goal is to shed new light on the rhetorics of radicalism, as well as the dynamics of their usage. To do so, we adopt a relational approach, paying particular attention to the interplay between themes in radicals’ writings and how this interplay is shaped by radicals’ social relationships. Specifically, we examine two questions. First, which, if any, rhetorical patterns help constitute radical discourse? Second, why do radical groups adopt one rhetorical pattern, or a mixture of patterns, instead of another?

We identify and conceptualize radicals’ rhetorics through a novel use of computational textual analysis in the task of theorizing. We call this approach *computational abductive analysis*. Abductive computational methodology “defamiliarizes” the concept under study to expose its puzzling aspects. Resolving these puzzles through an iterative movement between a widening body of empirical material and further analyses can help generate or refine conceptual frameworks, thereby laying the foundation for new theories (Tavory and Timmermans 2014).

Our computational abductive analysis builds on the radicalism scholarship, our knowledge of the empirical case, and a computational analysis of an original corpus of more than 23,000 pages published by radical groups operating in and around Afghanistan between 1979 and 2001. To the best of our knowledge, this corpus, comprising documents we collected in Afghanistan, Iran, and Pakistan and introduce here, is the

most extensive dataset of radical groups’ writings compatible with computational methodologies.¹

Previewing our findings, we first identify two ideal types of radical rhetoric. A *rhetoric of subversion* evokes an ongoing, historical-scale fight against powerful enemies external to radicals’ immediate community. In contrast, a *rhetoric of reversion* turns the rhetorical gaze inward, drawing on notions of everyday morality in the personal and local spheres to uproot undesired elements from among radicals’ close social context. This kind of rhetoric reveals an intimate form of radicalism—a kind often overlooked by common conceptualizations of radicalism as being pitted against despised outsiders. The second part of the article examines why radical groups adopt more of one rhetorical type than another. Drawing on insights into the interdependency of relationships, social domains, and discourse (*e.g.*, Bearman 1993; White 1995; Mische and White 1998; White 2008; Mische 2011), the analysis offers evidence that groups’ ties to supporters affect which rhetoric they deploy.

We conclude with a discussion of our study’s contributions to the scholarship on radicalism and, more generally, discourses of contention (*e.g.*, nationalism, populism). Specifically, we highlight how rhetorics of subversion and reversion both construct a discursive boundary of futurity (Mische 2009; 2014): some prevailing and powerful ideas, ways of life, individuals, groups, and institutions are rhetorically cast as not belonging in the future. Conceptualizing other discursive boundaries of exclusion can, in turn, help advance our theorization of contentious discourse, including ways these discourses split, overlap, and merge. As a whole, our study begins to chart a relational understanding of radical discourse by examining how relationships and boundaries affect the development and adoption of radical concepts, ideas, and language.

2 Patterns of radical rhetoric

Social scientists seeking to discover unknown concepts and themes in large textual corpora increasingly use unsupervised machine learning techniques—most commonly, topic models—to identify patterns of word usage (*e.g.*, Quinn, Monroe, Colaresi, Crespin, and Radev 2010; Grimmer and Stewart 2013; Miller 2013; Mohr and Bogdanov 2013; Roberts, Stewart, Tingley, Lucas, Leder-Luis, Kushner Gadarian, Albertson, and Rand 2014; Light and Odden 2017). This inductive analysis can be combined with expert knowledge, a deep engagement with source material, and other automated content analysis techniques to yield novel and compelling inferences, an approach recently described as computational grounded theorizing (Nelson 2017).

However, an inductive approach is less applicable when researchers come to the data already having some idea of the direction in which hypotheses, arguments, and theories may develop. When ethnographers find themselves in such circumstances, they often undertake abductive analysis: identifying surprising findings in light of prevailing arguments, and reasoning through what makes them unexpected to generate a newly refined theory with greater explanatory power (Tavory and Timmermans 2014). Similar to ethnographers familiar with arguments relevant to their field site, we have a wealth of scholarship on radicalism to draw

¹ For a similar kind of corpus consisting of writings by individual Arabic-speaking jihadists, see Nielsen 2017.

from, as well as literature on and experience with our empirical case, Afghanistan. As a result, we develop an abductive framework for the analysis of a large corpus of radicals’ writing to conceptualize the rhetorical underpinnings of radical discourse.²

2.1 Afghanistan, crucible of radicalism

Our study relies primarily on two decades of text produced by radical groups in and around Afghanistan.³ This period of time begins with an uprising against the central government in 1979, stretches through a long and evolving conflict, and ends on September 11, 2001. During this period, radical groups used printed text as a key medium for broadcasting their rhetoric to diverse audiences at home and abroad⁴—a necessary task due to the country’s entanglement with regional and international power-struggles. For example, magazines, newspapers, speeches, sermons, and pamphlets could be printed in multiple languages to entertain local and foreign fighters in the battlefield (Rana 2008), sent abroad to attract financial support and new recruits (Rubin 2002), and distributed in refugee camps and on university campuses—where the early radical groups formed (Roy 1986)—to boost support and draw recruits, a tactic still used today.⁵

Studying the period between 1979 and 2001 affords two analytic advantages for examining radical rhetoric and its adoption.⁶ First, there is variation in rhetoric, both across and within radical groups. The numerous groups operating in and around Afghanistan during this time espoused a range of ideas and arguments that formed a general discourse of radicalism, and some of these groups changed their rhetoric over time (Christia 2012). Second, these groups differed in ways that could influence their generation and adoption of rhetoric. For example, some *mujahideen*, the rebels who fought Afghan communist and Soviet forces during the Jihad (1979 to 1989), received more financial and military support from the United States and its allies than other *mujahideen*. Similarly, some were more active on the battlefield, and suffered more casualties, than others (Rashid 2000; Rubin 2002). We collect data on these relevant group characteristics for use during the second part of our study.

In addition, this period in Afghanistan’s history helps us understand contemporary radicalism. Not all radical discourse and behavior is tied to Islam, but there has been an increase in activity by militant Muslim groups over the past few decades (Walter 2017b). Many of these groups’ lineages can be traced back to Afghanistan, 1979 to 2001. For example, the Islamic State arose from al-Qaeda of Iraq, an affiliate of the original al-Qaeda, which was organized during the Jihad and has roots in Abdullah Azzams *Maktab al-*

² Others have noted the affinity between the logic of abduction and both data science (Kitchin 2014; McFarland et al. 2016) and comparative-historical sociology (Anderson 2018). We parallel these efforts by describing an abductive computational analysis of textual data. We thank the *ASR* editors for noticing the applicability of abductive inference to our study.

³ See Appendix A for a discussion on the selection of these groups and the data collection procedures.

⁴ Another important medium was speech, such as sermons, which continue to play an important role in spreading jihadist ideas (Nielsen 2017). However, radical groups printed materials were a crucial part of communication strategies because they could be spread widely. A count of publications printed by Afghan, Pakistani, and transnational jihadist groups based in Pakistan in 1990 tallied 100 monthly magazines and 12 weekly newspapers (Rana 2008).

⁵ See <https://tribune.com.pk/story/797526/islamic-state-pamphlets-distributed-in-south-waziristan/>, last accessed 28 November 2018 and <https://www.pajhwok.com/en/2017/03/12/leaflets-call-logar-youth-join-daesh>, last accessed 12 December 2018.

⁶ We use September 2001 as an endpoint because the subsequent invasion by the United States and its allies dramatically reoriented discourse in Afghanistan. The rhetorics we conceptualize could have continued after September 2001, but this date provides a significant moment of discontinuity that justifies a closing bound to our data collection.

Khidamat (the Services Bureau), one of the groups included in our study. Thus, our interest in Afghanistan extends beyond its analytic usefulness—we gain a window into the early formulations of one of today’s most prominent radical traditions.

2.2 Textual data

We built our corpus by collecting magazines, newspapers, speeches, and pamphlets in Afghanistan, Iran, and Pakistan. These materials were originally produced by 22 different individuals or groups operating in or around Afghanistan between 1979 and 2001.⁷ The authors include both famous and lesser-known *mujahideen* organizations—for example, *Jam’iyyat-i Islami* and *Hizb-i Islami* (Gulbuddin faction), in contrast to *Jami’ah al Da’wah ila al-Qur’an wa al-Sunnah* and *Hizbullah Afghanistan*—as well as the eventual enemy of some *mujahideen* groups, the Taliban. We found relevant material in five languages: Arabic, Dari, English, Pashto, and Urdu. Appendix A provides details of our data collection and text preparation.

The corpus has three characteristics allowing us to uncover meaningful rhetorical patterns within radical discourse, and, later, the reasons for their adoption. First, our corpus captures as wide a swath of the period’s radical discourse as possible. Of course, as with nearly all datasets of historical material, it does not contain the entire population. However, we approximate that population by including texts from a wide range of groups, origins, and languages. For example, instead of only including publications from large *mujahideen* groups, we sought out texts produced by more obscure groups. Similarly, rather than focusing solely on the famous Sunni-majority “Peshawar Seven” groups, we include publications by Shi’ite-dominated groups. In addition, we collected publications by the same group in different languages whenever available. Furthermore, our corpus includes full runs or near-full runs of many publications, including those produced by influential groups like *Jam’iyyat-i Islami*, the Taliban, and precursors to al-Qaeda. In summary, we are confident that our extensive data collection has minimized, and resulted in non-systematic, omissions. For more information on the coverage of our corpus, see Appendix B.

A second advantage is that the data do not suffer from the response errors of social desirability bias and inaccurate recall common in survey and interview data. The documents capture the rhetoric that radicals chose to use. Finally, the corpus offers observations over many years. We exploit this longevity in the second part of our study, when we use variation in global oil prices over time to identify the direction of the relationship between networks of support and rhetoric.

2.3 Computational abductive analysis

We developed our conceptualization of radical rhetoric via a computational abductive analysis of textual data, presented in the following section. This kind of analysis parallels work using ethnographic and comparative-historical data (Tavory and Timmermans 2014; Anderson 2018), except we rely on computational method-

⁷ We also acquired some publications from the Hoover Institution and the Taliban Sources Project (Strick van Linschoten and Kuehn 2018), but these materials were also originally gathered in Afghanistan.

ologies to “defamiliarize” the object, or concept, under study.

We began as other abductive analyses do: we studied existing social scientific theories and arguments regarding radicalism, as well as the details of the case. The latter involved multiple rounds of informal interviews in different parts of Afghanistan, reading the secondary literature, and a close reading of two independent journalistic publications that were collected alongside the radical texts but not included in the corpus.⁸ Both of these sources, *Afghanistan Forum* (1983 to 1985) and *Afghanistan Information Center Bulletin* (1988 to 1992), were sympathetic to the anti-Soviet cause but were not allied with any particular militant group. During this initial stage, we surmised that radical discourse is commonly understood as directed outward: radicals attack powerful individuals, organizations, ideas, and ways of life in their society, but outside of their immediate community and network of supporters.

Next, we identified latent categories of meaning, or “topics”, running throughout the corpus using a structural topic model (Mohr and Bogdanov 2013). Topics are defined using frequency distributions of terms probabilistically estimated to cluster together; they are themselves distributed across documents in the corpus (Blei, Ng, and Jordan 2003; Grimmer and Stewart 2013). Put in the language of the sociological study of discourse, topic models first show us collections of words based on how they appear together. Then, for each topic, the words can be interpreted in relation to one another to gain insight into the meaning represented by the topic (Somers 1995; Mohr 1998; Somers 2008).

Identifying topics in the corpus facilitated a key methodological step: defamiliarizing radicalism. That is, we used the computational measurement of the corpus to identify puzzling aspects of radical expression.⁹ In our case, we found that although several topics corresponded with the scholarly literature’s prevailing understanding of radical rhetoric—what we call “rhetoric of subversion”—some did not fit. This compelled us to ask questions about radicalism that we otherwise would not have (Tavory and Timmermans 2014: 57). The pursuit of these questions led to a recursive movement between the computational results, a close reading of selected corpus material, further literature on social movements, and additional theories that engendered a novel conceptualization of radical rhetorics. The computational abductive approach also involved systematically iterating through a range of topic model specifications (see Wilkerson and Casas 2017).¹⁰

The first step in our computational examination was to prepare the text data in line with recently developed strategies for comparative computational textual analysis. Non-English texts were translated into English (Lucas, Nielsen, Roberts, Stewart, and Tingley 2015), all the text material was preprocessed, and publications were parsed into documents, or pages of the publications, that act as our observational units (Grimmer and Stewart 2013). This preparation yielded 23,343 documents (see Table 1).¹¹ Next, we

⁸ The journalistic sources were found serendipitously during collection of the radical texts. We were not previously aware of their existence and could not a priori assess their usefulness. As a result, we read them in their entirety.

⁹ Defamiliarization resembles the “decontextualization” that takes place in computational grounded theory. But, whereas decontextualization refers to the simplification of text into patterns that were not initially noticeable (Nelson 2017), defamiliarization casts part of the object as a problem that requires a solution drawn from theory generation (Tavory and Timmermans 2014).

¹⁰ For other techniques with the potential to defamiliarize an object in textual data, see Breiger et al. 2018.

¹¹ For details and a discussion of our preprocessing, see Appendix C.

estimated a structural topic model (STM). An STM works by incorporating metadata, or information about documents, as covariates (Roberts, Stewart, and Airoldi 2016). We included the covariates of authorship (*i.e.*, the individual or group that published the text), year of publication, and the original language of the text.¹² This way, we isolated patterns in the semantic content of the corpus rather than differences due to idiosyncrasies related to any values of the covariates (Lucas et al. 2015). After an iterative process evaluating model quality based on qualitative validation (Grimmer and Stewart 2013), we selected a 10-topic solution. This solution was further validated using a data-driven assessment. A detailed discussion of our computational procedures, including the selection of the number of topics, validation assessments, and robustness checks can be found in Appendix C.

We then drew on the model’s output of topics, a close reading of documents highly associated with specific topics, and our knowledge of the case to understand each topic. Following convention, we labeled the topics based on our understanding; this helps make them more accessible to readers unfamiliar with the case (Mohr and Bogdanov 2013). When examining topics’ words, we paid special attention to their FREX scores, a measure of terms’ probability of appearing under a topic and their exclusivity to that topic (Roberts et al. 2014). Next, we interpreted them individually and in groups in light of the prevailing ideas of radical discourse. When we encountered topics that challenged these ideas, we turned to a broader literature—specifically, research on religious social movements—to construct a new conceptual framework of radical rhetoric. Our resulting conceptualization identifies and characterizes two rhetorics of radicalism.

Of course, textual material—whether entire publications or individual pages—contain a mixture of rhetorical types. Therefore, in the final step of the analysis, we calculated the prevalence of each topic per document and used the topics’ associations with one of the two rhetorical types to determine a *rhetoric ratio score* for each document. This score indicates the relative prominence of rhetorical types at the document level. The more a document is made up of (multiple) topics forming one rhetorical type, the more its score approaches one or zero. The rhetoric ratio score can be used to index the rhetorical variation throughout the corpus, as well as further check our interpretation. Namely, previously un-read documents selected solely by an extreme rhetoric ratio score should contain text closely aligning with one rhetorical type or the other.

Finally, we used our conceptualization of radical rhetoric to specify two general dimensions structuring radical discourse, connectivity and reach, or how radicals are urged to achieve their goals and on what timescale (Mische 2009). Where radical rhetoric falls on these dimensions gives it its character. In the Discussion section, we explore how considering rhetorical connectivity and reach helps shed light on other forms of contentious discourse.

¹²After completing the main analysis, we re-estimated the STM with the inclusion of publication type (*e.g.*, magazine, newspaper, speech) as a covariate. There was no significant change in the topic content. See Appendix C.

Table 1: Corpus Description

Feature	Corpus
Number of authors	22
Publication locations	Afghanistan, Iran, Pakistan
Span of publication dates	September 1979 through September 2001
Unique publication titles	72
Individual publications	990
Documents (N)	23,343

2.4 Varieties of radical rhetoric

Radicals are typically defined as individuals or movements undertaking extra-institutional opposition against an external, and often hegemonic, target (Ferree 2003; Beck 2015). This opposition may be nonviolent or violent (Dalgaard-Nielsen 2010; Bartlett and Miller 2012; della Porta 2018), but it is a struggle directed toward (contextually) powerful individuals, organizations, ideologies, and sociopolitical and economic systems radicals understand as extrinsic. For example, Calhoun (2012: 6, 31) labels labor movements in nineteenth-century Europe as radical once they became “deeply at odds with prevailing conditions . . . and the [industrializing] policies of dominant elites.” Similarly, Marks, Mbaye, and Kim (2009) differentiate radical leftist parties in early twentieth-century Europe from their mainstream counterparts by their opposition to systems they wanted no part of (*i.e.*, capitalism and private property). Animal activists are deemed radical when they harm companies abusing animals (Loadenthal 2013) and, as Hegghammer (2009) notes, scholars often perceive Salafist Muslims as radical once they adopt an anti-Western stance (see also Roy 2017: 41). In the context of the United States’ abortion-rights movement, radical feminists, Ferree (2003) argues, are those who challenge the prevailing pro-choice discourse with a “feminist pro-life” position. How do our data fit with this understanding of radicalism?

2.4.1 Radicalism of subversion

In our corpus, we uncovered extensive evidence in support of conceptualizing radical thought and speech as externally oriented. The STM, for example, yielded topic 1, “war”, and topic 5, “Afghan jihad (battle)”, both of which capture the notion of fighting external opponents, such as enemy soldiers (topic 1) and the Soviets and Najibullah Ahmadzaib (topic 5), “Najib”, the leader of the communist regime in Kabul (see Figure 1). Indeed, the terms “Afghan” and “jihad” in topic 5 reflect the concept of the “lesser jihad”, or the armed fight against the enemies of Muslims (Nielsen 2017: 107), giving the topic its name.

Similarly, topic 7, “international politics”, and topic 8, “social and political revolution”, regard the international and broad sociopolitical arenas—realms external to the radicals—as sites for struggle. Topic 7, for example, highlights the countries involved in contemporary Afghanistan’s long conflict, Russia and the United States. Topic 3, “theology of jihad (struggle)”, has much in common with topics 1, 5, 7, and

Figure 1: Labels and Top FREX Words for Selected Topics

1.	War	mujahideen, kill, attack, soldier, capture, enemy, operation, command, tank, area
3.	Theology of jihad (struggle)	god, jihad, quote, messenger, allah, brother, prophet, sheikh, bless, martyr
4.	Local education & markets	education, thousand, province, news, train, respect, ministry, school, abdul, mullah
5.	Afghan jihad (battle)	mujahideen, jihad, soviet, najib, kabul, regime, afghan, election, president, pakistan
6.	Religious code	movement, holy, life, great, human, scholar, woman, quran, answer, world
7.	International politics	afghanistan, russian, countries, govern, unit, international, russia, support, state, american
8.	Social & political revolution	social, unity, political, communities, revolution, parties, west, exit, nature, system

Note: Endings have been added to stemmed terms based on the usage of words in each topic's highly associated documents. For the full STM results, see Appendix C.

8: it references the religious dimension of jihad—the “greater jihad”, or an ongoing quest for godliness (Nielsen 2017:107)—rather than physical and political battles, but it does so by invoking a continuous fight against external enemies. Highly associated passages contain abstract exegeses on theological matters and hagiography combined with depictions of a virtuous religious community's resistance to threatening, impious outsiders.

Note that topics 1, 3, 5, 7, and 8, as a set, capture language that often uses historical references to describe the ongoing battle, underscoring how enemies have plagued radicals' historical community for centuries. For example, strongly associated documents contain passages such as the following:

The [Catholic] Church's most dangerous crime was its lying, misleading propaganda against Islam throughout the Crusades and after them, and its portrayal of Islam as a backwards, pagan

religion. This filled Europeans with a strong hatred and violent aversion to Islam, a hatred and aversion that have been passed down, generation to generation ... When Europe rebelled against the illogical teachings of the Church, it reverted to its pagan origins, giving new life to its ancient Roman and Greek traditions that were built on materialistic atheism ... It was atheist in everything, except in its dealings with Muslims. With the Muslims, Europe is a Christian crusader in which the secular state allies with the Church, and men play the role of the evangelizing monk, the orientalist scholar, and the professional spy, all at the same time. In conclusion, we present the words of the English tyrant [William] Gladstone (referring to the Quran): “So long as there is this book, there will be no stability in the East.”¹³

Muslims, everywhere and at every point, have suffered pain and troubles because of their beliefs, and still continue to do so. For instance, Muslims were horrifically mass murdered in Spain only because they were Muslims. Today in Afghanistan, Muslims are being mass murdered at the hands of Russian beasts because of this very reason. The Afghan people are, as a collective, Muslims and they are not ready to leave their faith.¹⁴

Thus, the first set of topics forms a rhetorical pattern consistent with common understandings of radicalism. Together, they depict a group—the radicals—attacking external opponents, whether actors or ideas. The depiction uses allusions to history, communal identities, and violent struggle to vilify and undermine well-established external hegemonic ideas and powerful actors, or perceptions of those ideas and actors. We term this kind of rhetoric a *radicalism of subversion*.

2.4.2 Radicalism of reversion

What about the other topics? Topic 4, “local education and markets”, captures radicals’ interest in everyday private, familial, and local matters, such as educational instruction and how one should treat various illnesses. For example, the thousand in this topic refers to prices—“ten thousand afghanis”—being set for local markets. Topic 6, “religious code”, is made up of religiously justified regulations of everyday life. Many passages associated with this topic are *fatwas*, or advisory legal opinions commonly used in Muslim communities. *Fatwas* may deal with macro-level sociopolitical conflict (Nielsen 2017), but the texts linked to topic 6 focus on the everyday behavior of individuals. We found that the remaining topics are not relevant to general sociopolitical discourse, and we do not consider them further in our analysis.¹⁵

Topics 4 and 6 do not invoke an ongoing, historical-scale offensive against external enemies, raising questions about how they fit with the predominant understanding of radicalism. Indeed, documents highly associated with these topics depict a distinctly intimate concern, centering on radicals’ own lives and com-

¹³From *Al Mujahid*, by *Jami’ah al Da’wah ila al-Quir’an wa al-Sunnah*, March 1989.

¹⁴From *Mashal*, by *Jam’iyyat-i Islami*, September 1986.

¹⁵It is not uncommon to uncover non-relevant topics since it is not necessary that all words in a corpus relate to the research question. In addition, some topics may be formed as “residual categories”, or collections of words defined by their lack of fit in the other topics. For further discussion on omitting the remaining the topics, and for the full STM results, see Appendix C.

munities. For example, one document includes an opportunity to criticize Western science, but instead the focus is on guiding the behavior of individuals within the community, such as local religious figures:

Question. Some *imam sahebs* [religious authorities] are saying that science is the knowledge of the devil. Answer. This is a wrong assertion they hold. If you read the Quran it says prepare yourselves from the unbelievers for jihad to the extent that you can. The unbelievers are those who have planes, canons, machinery, Kalashnikovs, and other technical equipment and are making them. Does it mean that Muslims should not be making such technical things? The fact is that respected *imams* and all the rest should learn these professions and knowledge... The *mullah* [religious leader] will not hold such an opinion.¹⁶

To help make sense of this unexpected language, we engaged a broader range of literature, deriving especially useful insights from the scholarship on religious social movements.¹⁷ This work observes that a number of religious social movements, ranging from a proselytizing Buddhist group (Snow, Zurcher, and Ekland-Olson 1980) to early Protestants (Goldman and Pfaff 2018), have sought to effect social change by foregrounding struggles internal to individuals and communities, often by fixating on the intimate spheres of gender, sexual behavior, and the family.

The religiously motivated temperance and anti-slavery movements in nineteenth-century America provide an instructive example. Rather than prioritizing attacks on the state and political parties, these movements worked to transform the United States by encouraging affiliated church communities to instill pledges and confessions of sinful behavior. Each individual was compelled to “deny himself, [and] abandon his prejudices and error, and become a firm, active, and a pledged friend of” the movements (the New York State Temperance Society 1833; quoted in Young 2002).

The strategies of these religious social movements point toward a new understanding of topics 4 and 6. Namely, these topics indicate how radicals can, at times, adopt an internal gaze. The language turns inward—spatially and temporally—toward individuals, families, communities, and contemporary concerns. Behavior and beliefs are judged and policed with little historical context, relative to topics 1, 3, 5, 7, and 8. This is often associated with religious doctrine, as in our case and in the cases of U.S. temperance and abolitionist movements (see also Friedland 2002; Brubaker 2015), but it is not necessarily religious. Extreme right-wing political movements in the United States, for example, have first adopted a moralizing language and then subsequently argued they were advancing “Christian” goals (Lipset and Raab 1978).

Of course, the features of intimacy and immediacy do not make language radical. Rather, we interpret topics 4 and 6 as jointly capturing a variant of radical rhetoric because they use intimacy and immediacy to construct an everyday, personal, and moralizing framework for fundamentally changing society. That is, together they advocate for the transformation—and, sometimes, destruction—of individuals and communities for the purpose of uprooting transgressive and unwanted ideas, traits, and lifestyles. We see this framework

¹⁶From *Tolo-ye-Afghanistan*, by the Taliban, October 1996.

¹⁷We thank an anonymous reviewer to suggesting the literature on religious social movements.

implemented in a text prominently featuring an interview with a captured (alleged) spy, who, after time spent in captivity, says,

In the beginning [of my captivity], I was hoping that [the guards] might send me back to my family. But now I have been born again. . . I am ashamed, ashamed of what I have done to the brothers here [the radical group that produced the publication]. . . So I ask them. . . if it is possible that they can send me to Jihad. I am 22 years old now and I have spent those 22 years in disobedience to Allah. So. . . I ask Allah that he make the brothers a factor in my being recruited for Jihad in the path of Allah.¹⁸

This vivid portrayal of self-denial and personal transformation ends with the spy being asked if a death sentence is appropriate for his transgression; he agrees that it is.

In contrast to a radicalism of subversion, the rhetoric composed of topics 4 and 6 adopts a sociotemporally intimate gaze, stitching together concepts of morality, the everyday, and the personal and local to challenge members of radicals' own groups and communities, as well as the radicals themselves. Because this rhetoric is largely characterized by its inward turn, we label it a *radicalism of reversion*. Radicalism, in sum, is not only about the object, but also about a return, or reversion, to the subject.

2.4.3 Rhetorical mixtures

Rhetorical mixtures. Two varieties of radical rhetoric have emerged from our computational abductive analysis: a rhetoric of subversion and a rhetoric of reversion. Yet, to be clear, there are likely other rhetoric patterns embedded in radical discourse.¹⁹ In addition, our two variants are ideal types. Radicals' discourse and writings will contain a mixture of the two.

To explore the blending of rhetorics, we constructed a rhetoric ratio score based on the proportion of each topic in the document. The more a document is made up of the topics underlying the rhetoric of subversion, the more that document's rhetoric ratio score approaches one. The greater the proportion of topics that constitute the rhetoric of reversion, the closer a document's score is to zero. Documents with a more even mixture of rhetorical types receive a score closer to 0.5.

With documents placed along this continuum, we can help verify our analysis by selecting a subset of documents based solely on their score, then comparing their contents to our theorized rhetorical types. The example passages in the upper panel of Figure 2, with rhetoric ratio scores of 0.90 and 0.78, illustrate how a radicalism of subversion uses historically evocative language to attack ideologies and persons external to the radicals: nationalism and impure Others. We also see how the latter example's slightly lower score reflects its incorporation of notions of behavior and personal morality—elements common to a radicalism of reversion. Stronger examples of the reversion rhetoric are presented in the bottom panel of Figure 2. These passages,

¹⁸From *The Islamic Emirate*, by the Taliban, July 2000.

¹⁹We encourage future research to build more expansive corpora, which will likely help uncover and conceptualize further radical rhetorics.

with scores of 0.29 and 0.33, show how this rhetoric uses the language of everyday, micro-level behavior and morality, largely divorced from an ongoing, historical struggle against ideas and outsiders.

2.4.4 Dimensions of radical rhetoric

We conclude the identification and interpretation of rhetorical varieties by suggesting two dimensions that characterize radical rhetoric: *connectivity* and *reach* (see Mische 2009).²⁰ The first, connectivity, captures how the rhetoric urges radicals to achieve their goal, ranging, in our case, from externally directed attacks against despised outsiders and ideas to inwardly directed calls for personal and local transformation. It encompasses the imagined causal links moving radicals toward their desired outcome (Mische 2014: 452). The second, reach, expresses the timescale for achieving the sought-after societal transformation. Is radicals' desired society unfolding in the context of a long, historical struggle, stretching from the distant past into a far horizon, or is it manifested through an immediate uprooting of unwanted aspects of self and community?

The dimensions of connective and reach offer an initial infrastructure for theorizing radical discourse. If radical discourse consists of rhetorics defined by connectivity and reach, then it helps construct a specific projection of the future. After all, these dimensions articulate, first, a logic of how to arrive at an imagined future, and second, a timescale for arriving there (Mische 2009; 2014). We will build on this insight into radicalism's *futurity* and use it in comparison with other discourses of contention to propose first steps toward a theory of radical discourse. First, however, we extend our understanding of radical rhetorics by examining the reasons for their adoption.

3 Adopting rhetorics

What leads radicals to embrace one variety of rhetoric over another? Why would a radical group espouse a radicalism of subversion over one of reversion? We propose an explanatory argument of rhetorical development and adoption that builds on research on violent political insurgents and the role of social relationships in discourse. We tie together these literatures with the simple premise that discourses of contention, such as radicalism and populism, are dynamic interactional processes (Bonikowski and Gidron 2016). In other words, ongoing social relations among actors sharing a social context drive the development of meaning, ideas, and rhetoric (Snow, Rochford, Worden, and Benford 1986; Mische 2003; Snow, Benford, McCammon, Hewitt, and Fitzgerald 2014).²¹ For example, social connections helped determine whether rebels adopted localized or abstract religious identity discourses during the run-up to the English civil war (Bearman 1993), and stronger local ties contributed to protesters' notions of neighborhood solidarity during the Paris Commune (Gould 1995).

²⁰ Conceptualizing two dimensions suggests there are at least four kinds of rhetorics of radicalism. As noted earlier, our analysis yielded two, and we encourage future research to explore other potential rhetorics.

²¹ Or, as Mische and White (1998: 695) put it, "discourse is the stuff of social networks."

Figure 2: Examples of Varieties of Radical Rhetoric

Radicalism of subversion

During the Battle of Uhud [CE 624], a man named Quzman fought violently against the pagans, to the point that he himself killed seven or eight of the pagans. After the end of the battle, the companions found him, wounded, and carried him to the house of Bani Zafr, where they comforted him. He said, “By God, I only fought for my people, and if it were not for them I would not have fought.” When the pain of the wound intensified, he killed himself. Whenever his name was mentioned, the Prophet, prayers and peace be upon him, said, “He is in hell.” This means that whoever fights because of patriotism or nationalism, or for any reason other than elevating the word of God, this is his fate, even if he fights under the banner of Islam. [From *Al Mujahid*, by *Jami’ah al Da’wah ila al-Quir’an wa al-Sunnah*, May 1989. **Rhetoric ratio score = .90**]

Preparation can only be achieved through pure and clean people, people who are clean inside and outside, in private and in public, in heart and in form. These people are only found in the available appropriate places, not in the places where the true believers refuse to look. Believers do not seek the places that focus on worldly finery and pleasures, and are concerned with this life ... Preparation can only be achieved with pure content and material, in which the solution truly lies. This will bring victory, and, at its forefront, salvation and peace ... It is chosen for a group of good people who emigrated following the path of God, and have prepared to continue on the path forwards towards the better and the more comprehensive, which will never result in falsehood. [From *Source of Jihad*, by Miramshah College, November 1991. **Rhetoric ratio score = .78**]

Radicalism of reversion

Question. Is food cooked by unbelievers ok for Muslims to eat? Answer. There are two kinds of unbelievers. Those who believe in one of the heavenly books like the English or the Americans; if they cook food it is fine for the Muslim to eat it. But food cooked by atheists like the Russians or Hindus is not kosher for Muslims. However, their women are *halal* [religiously clean] for Muslims to marry, but if they cook food and meat with vegetables like pulses, the Muslim can only eat the vegetables and the pulses from it. [From *Tolo-ye-Afghanistan*, by the Taliban, October 1997. **Rhetoric ratio score = .29**]

It should be clearly understood that every pious and faithful Muslim believes firmly that divorce is undoubtedly the most unwanted and undesired situation and one should not have a recourse to it except when all the other doors for reconciliation are closed and mediation and arbitration by relatives and friends do not prove fruitful and domestic life has been miserable ... After the divorce, Islam contains clear instructions for the protection and upbringing of their children and commands the husband to provide financial help to the minor children, so long as they attain adulthood and become independent. In the meantime, he has to provide financial help to the divorced spouse for a certain period according to the rules of the Islamic Code. Even during separation, Islam enjoins respectful behavior towards each other. [From *Afghan Jihad*, by Cultural Council of Afghanistan Resistance, December 1992. **Rhetoric ratio score = .33**]

3.1 Argument and hypothesis

We argue radicals relationships of material and social support affect which variety of rhetoric they adopt. This occurs because of two mechanisms. First, the ties through which groups access resources strengthen certain social relationships over others. Namely, when receiving support from extra-local or foreign patrons, groups tend to have weaker social ties with the local populace (Salehyan, Siroky, and Wood 2014; Beardsley, Gleditsch, and Lo 2015). For example, when rebels weigh whether to use terrorist tactics, which tend to undermine their legitimacy among locals, those relying on non-domestic sources of support are more likely to commit acts of terrorism (Fortana, Lotito, and Rubin 2018). This consequence of external support was well known during the Afghan Jihad. Rubin Rubin (2002: 181) notes that “as aid of all sorts increased, [radical groups] became less dependent on their local population. Increasingly, they lived in bases (*qarargha*) rather than in their homes and villages.”²²

Second, the direction in which radicals’ relationships are strengthened then influences their rhetoric. For instance, in 1988, a member of an Afghan radical group known locally as the “Wahhabis” because of its affiliation with Arab states along the Persian Gulf, where Saudi Arabia has long promoted the Islamic doctrine of Wahhabism, announced from a prominent mosque in Afghanistans Kunar province that “Kunar is not fully free. Kunar will not be free unless we conduct purges and eliminate nationalist elements and tribal relationships.”²³ The group’s goal, of course, was to replace a local or national governing logic with a rule of law based on Wahhabism. Similarly, the Afghanistan Information Center, an independent journalist agency, observed in 1991 that many of the groups with strong backing from Islamists in the Arab Middle East and Pakistan were aiming for “the imposition of revolutionary Islam, [and undertaking] extensive efforts to destroy traditional systems in the country and to replace them with an artificial party system [that will] further disappoint the people.”²⁴

Different social processes can explain why radicals’ relationships of support affect the kind of rhetoric they adopt. The scholarship on movement framing suggests radicals craft their rhetoric over time to converge with the beliefs and values of (potential) supporters, in part to facilitate the continued flow of resources (Snow et al. 1986; Snow and Benford 1988; Snow et al. 2014). Social exchange theory, which contends that exchange relationships confer benefits to both parties (Blau [1964] 2008; Molm and Cook 1995; Cook, Cheshire, and Gerbasi 2006), implies that supporters of radicals should expect something in return for providing resources, such as the promotion of their beliefs and values. Consequently, radical groups are likely to adopt rhetoric reflective of their backers’ ideology.

However, radicals primary goal is to seek fundamental change (Beck 2015), and such change might entail altering the social, political, and economic structures enabling their backers to provide resources. As a result, although radical groups may often adopt a rhetoric that aligns with the perspectives of supporters, they could

²²The sociological research on brokerage provides complementary evidence: brokers can be swayed toward forming stronger bonds with the more influential or powerful side of the brokering relationship (Stovel and Shaw 2012).

²³From *Afghanistan Information Center Bulletin*, October 1988.

²⁴From *Afghanistan Information Center Bulletin*, August 1991.

also come to contradict their backers (see Ferree 2003). For example, in the case of Afghanistan, *Hizb-i Islami* (Gulbuddin faction) received much support from Pakistan, which itself was supplied by the United States, yet its publications often contain claims such as, “the Americans and the Russians are united against Islam” and “the United States is...hesitant at the prospect of a free Afghanistan, exposing their duality of character.”²⁵ Such potential for divergence is not easily explained by commonly proposed processes, including those suggested by framing and social exchange theories.

We posit a contingent process that unfolds as a radical group encounters its supporters’ social world. Before this encounter, radicals and supporters are ensconced in their respective domains, largely defined by the interplay of social ties, discursive signals, and stories and narratives (White 1995; Mische and White 1998; White 2008).²⁶ When the provision of resources occurs, and social relationships across domains emerge, the intersection of the radical group’s and the supporter’s domains engenders an “interstitial space” defined by their new co-presence (Mische and White 1998: 705). In this space, or “public”, values, identities, and narratives rooted in each domain are suppressed, and the actors commonly assume identities and perspectives that minimize potential tension during the encounter (Mische and White 1998; Mische 2007; 2011).

In our case, we expect that when foreign patronage pulls some radicals towards an extra-local domain, these radicals are more likely to espouse a rhetoric of subversion, in part by using tropes from the historical and long-term view that often imbues international patrons’ own outlook and strategy. The connectivity of their rhetoric will skew toward a macro-level assault on external enemies. The rhetoric’s reach will situate the radicals’ fight in an unfolding, historical conflict stretching into the horizon. In contrast, if support is predominantly localized, then radicals’ rhetoric is more likely to focus on local and individual behavior and morals. The temporal outlook, influenced by the primacy of local concerns, is constrained to the proximate past and near future. That is, the connectivity is intimate and the reach is close. To be clear, this mechanism explains the adoption of a type of discourse, not why radicals might come to believe supporters’ values and goals, or vice versa.

The notion of publics is key for understanding radicals’ adoption of rhetoric. It accounts for straightforward, likely convergent, encounters between existing domains, as well as unexpected contention. In other words, dynamics of publics underscore how connections between actors from different domains may help hold “rivalries, critiques, [and] resentments” in abeyance (Mische and White 1998: 708). Yet, they also explain how, in other cases, actors’ domains—their existing beliefs, views, and narratives—might come to the foreground, sparking contradictions, a reinterpretation of narratives, and clashes (Mische and White 1998; White 2008). For example, radicals may find their own domain of deep-rooted social change pushing against their supporter’s domain. This disruption might compel radicals to use the rhetoric that originally resonated with supporters to instead repudiate them. Thus, thinking of relationships engendering publics, and the navigation of existing domains that occurs within publics, allows for an independent and creative construction of rhetoric. Relationships define a public, within (and during) which radicals experiment, problem-solve,

²⁵ From *Mujahideen Monthly*, by *Hizb-i Islami* (Gulbuddin faction), April 1990.

²⁶ White (2008) calls these social worlds as “netdoms”, or network domains.

and craft rhetoric in light of their domains (Emirbayer 1997; Jansen 2017).

Following the two mechanisms we have detailed, we propose a hypothesis for why radicals adopt one variety of radical rhetoric over another:

H1. *Receiving external support increases the probability that radicals will espouse a rhetoric of subversion.*

The inverse of this hypothesis is that lower levels of external support will likely lead to greater connection with locals and, as a result, development of a radicalism of reversion.

3.2 Data on groups and their support

In our analysis of rhetorical adoption, the outcome variable is the document rhetoric ratio score. This variable takes a value between zero and one, indicating the document’s respective mix of radical rhetorics. The main predictor variable is whether or not radical groups received financial support from a foreign entity each year. This variable is measured dichotomously and drawn from the Uppsala Conflict Data Program (UCDP) External Support dataset (Högbladh, Pettersson, and Themnér 2011). From this dataset, we also know which countries provided the external support, information we use later in the analysis. We link groups’ external support, and their patrons, to their documents’ rhetorical scores across years via information on authorship and year of publication.

We also include data on other group characteristics that may potentially influence their discourse. Because the escalation of conflict has been found to affect radicalization (della Porta 2013; 2018) and rhetorical use (Isaacs 2016), we include the number of casualties a group suffers in a given year. We also include the number of fighters in each group over time, because a groups strength affects their ability to make and maintain alliances, which, in turn, influences the rhetoric they develop to justify the confederation (Christia 2012). Data for these two variables are taken from the UCDP Non-state Conflict dataset (Sundberg, Eck, and Kreutz 2012).²⁷ Table 2 presents the list of groups included in the external support analysis and summary statistics of the variables.

The primary limitation of these data is that not all 22 authors of our documents are included in the UCDP dataset. Fortunately, however, the overlap has good representativeness. The dataset contains most of the major groups, including two of the most influential, *Jam’iyyat-i Islami* and *Hizb-i Islami* (Gulbuddin faction); so-called moderate groups like *Jabha-yi Nijat-i Milli-yi Afghanistan* and *Mahaz-e Milli-ye Islami*; a major Shi’a group, *Hizb-i Wahdat*; and the Taliban. In addition, the missing groups do not differ in a significant way from those that are included with regard to their rhetoric ratio score. The mean rhetoric

²⁷We also consider groups’ ethnic and sectarian affiliation, which are well documented in historical accounts such as Rubin 2002 and datasets like ACD2EPR (<https://icr.ethz.ch/data/epr/acd2epr/>; Wucherpfennig, Metternich, Cederman, and Gleditsch 2012; Vogt, Bormann, Rüegger, Cederman, Philipp, and Girardin 2015). Ethnic and sectarian cultural traditions may constrain which historical tropes are available for inclusion in discourse (Jansen 2007). However, we ultimately do not include variables capturing ethnicity and sect in our models because there is little to no variation in rhetoric ratio scores across ethnic and sectarian affiliation. We infer this from the mean rhetoric ratio scores, and their standard deviations, shown in Table 2.

ratio score of the included groups (.66) is within a standard deviation of the missing groups' mean score (Table 2). Moreover, we reintroduce four groups in a second step of this analysis, which we conduct to shed light on causal direction.

Examining the direction of the association between support and rhetoric helps to address a second limitation of our observational data: an inability to unambiguously conclude that the effect of the main predictor is causal. However, we can provide evidence of causal direction using the exogenous moderating variable of monthly global oil prices, acquired from the International Monetary Funds public database.²⁸ We explain how we do so, as well as the overall analytic strategy, in the following section.

3.3 Analytical strategy

We specify a fractional logistic regression model (Papke and Wooldridge 1996; Wooldridge 2002) to examine the effect of a groups external financial support on their documents mixture of radical rhetorics.²⁹ Our initial model specification estimates the bivariate effect; the second specification estimates the effect after adjusting for group casualties and number of fighters.³⁰ Standard errors are clustered by group and time period.

The fractional logistic regression approach provides an estimate of the strength of association, not the direction of the association. As a result, the question of possible reverse causation remains: Is external aid driving rhetorical adoption, or do groups espousing a certain rhetoric receive more (or less) external support? To answer this question, we conduct a moderation analysis examining a conditional relationship between the origin of groups' support and global oil price. First, we divide the radical groups into two categories based on their foreign patrons. The "treated" category contains groups that received support from Saudi Arabia (SA) and other Arab states of the Persian Gulf region;³¹ the control category comprises groups without these SA or Arab Gulf connections. Then, we estimate a fractional logistic regression model using a moderator, the global price of oil, that we assume is unaffected by the relationship between external support and rhetoric (Brambor, Clark, and Golder 2006).

²⁸ The IMF reports oil price as a simple average of three price measures—Dated Brent, West Texas Intermediate, and the Dubai Fateh—in US Dollars.

²⁹ As Wooldridge (2002: 662) notes, an alternative to the logistic model would be to use a log-odds transformation of the dependent variable in a linear model, estimated by OLS. However, this approach cannot be used directly if the dependent variable takes on the boundary values, as our ratio measure could potentially do. Moreover, the parameters are difficult to interpret without further assumptions.

³⁰ The STM analysis accounts for the effects of group authorship, year, and language. However, as a robustness check, we include a specification that uses group fixed effects, as well as one including both group and temporal fixed effects (Appendix F). For the temporal fixed effect specification, we divide the years between 1979 and 2001 into five periods established in the literature on the Afghanistan conflict (Rubin 2002; Christia 2012). These periods are the early Jihad, the late Jihad, the intra-*mujahideen* civil war, the early Taliban period, and the late Taliban period.

³¹ In addition to those groups that are known to have received support from SA based on the UCDP dataset, four others are coded as having had strong ties to the Arab Gulf states because their membership included large numbers of individuals from these countries. These groups are the Haqqani Organization, Miramshah College, *Jami'ah al Da'wah ila al-Quir'an wa al-Sunnah*, and *Maktab al-Khidamat*. We test the assumption of affiliation based on membership in Appendix F.

Table 2: Summary Statistics of Data for External Support Analysis

Group Name	Documents	Mean Rhetoric Ratio	Mean External Support	Number of Years Receiving Support	Mean Casualties per Year	Number of Fighters	Ethnicity and Sect	External Supporters
Jam'iyyat-i Islami ^a	6,133	0.68 (0.17)	0.99	14 out of 15	121.8 (68.2)	12,000; 60,000 in 1996	Tajik, Uzbek; Sunni	US, Saudi Arabia, Pakistan
Jabha-yi Nijat-i Milli-yi ^b	1,839	0.69 (0.13)	1	All	NA	2,250	Tajik; Sunni	US, Saudi Arabia, Pakistan
Mahaz-i Milli-yi Islami-yi	611	0.63 (0.11)	1	All	10 (0)	2,500	Pashtun; Sufi, Sunni	US, Saudi Arabia, Pakistan
Ittihad-i Islami Bara-yi Azadi-yi	53	0.74 (0.08)	1	All	NA	1,250	Pashtun; Sunni	US, Saudi Arabia, Pakistan
Hizb-i Islami (Gulbuddin faction)	2,925	0.67 (0.13)	1	All	38.5 (35.0)	NA	Pashtun; Sunni	US, Saudi Arabia, Pakistan
Hizb-i Wahdat ^c	909	0.74 (0.12)	0.52	6 out of 8	21.3 (20.3)	116,000	Hazara; Shi'ite	Iran
Taliban ^d	2,066	0.48 (0.18)	0.48	3 out of 7	499.9 (127.8)	25,000	Pashtun; Sunni	Pakistan, Saudi Arabia
United Islamic Front for Salvation of Afghanistan ^e	609	0.62 (0.15)	0.77	2 out of 7	466.9 (327.4)	115,000	All; Sunni, Shi'ite	Russia, Iran
Other groups in corpus but not in UDCP dataset ^f	8,023	0.76 (0.18)	NA	NA	NA	NA	NA	NA

Sources: Högbladh et al. 2011; Sundberg et al. 2012.

Note: Standard deviation are in parentheses; mean external support indicates the amount of support received across all years; groups' number of fighters are not reported as a mean because of little to no variation across years; ethnicity and sect refer to the predominant affiliations (heterogeneity existed within some groups).

^aFor the external support analysis, *Jam'iyyat-i Islami Afghanistan* text data includes speeches from Burhanuddin Rabbani, its leader. ^bFor the external support analysis, *Jabha-yi Nijat-i Milli-yi Afghanistan* text data includes speeches from Sibghatullah Mojaddedi, its leader. ^cFor the external support analysis, *Hizb-i Wahdat* text data also includes text data from *Hizbullah Afghanistan* and *Islamic Dawat*, which joined other groups to form *Hizb-i Wahdat* in 1989. ^dTaliban text data includes data from the Taliban Support Council, a group based in London. ^eText data for the United Islamic Front for Salvation of Afghanistan (UIFSA), popularly known as the Northern Alliance, contains text data from *Jam'yyat* after 1999, when *Jam'yyat* began anchoring the UIFSA. The analysis includes data from 1979 to September 10, 2001, so US support for the UIFSA, which began after September 11, 2001, is not included. ^fGroups in the corpus but not in the UDCP dataset are excluded from the analysis of external support. An exception to this exclusion is the use of text data from the Haqqani Organization, Miramshah College, *Jami'ah al Da'wah ila al-Quir'an wa al-Sunnah*, and *Maktab al-Khidamat* in the moderation model.

The logic of this moderation analysis is as follows. Some radical groups’ external support originated with the governments and citizens of the oil-rich Arab states of the Persian Gulf, including SA.³² This support would have been tied to the price of oil. The higher the price of oil, the more monetary resources in SA or Arab Gulf states, and the greater amount of financial support flowing to some radical groups in Afghanistan.³³ In turn, we posit that more support intensifies radicals’ encounter with backers and compels radicals to further suppress their own domain in favor of their backer’s domain. As a result, for groups backed by residents and governments of oil-rich states, we would observe a stronger effect of external resources on the prevalence of subversion rhetoric when oil prices are higher, relative to groups without these backers. Because global oil price is exogenous to our proposed causal link, evidence of this effect would support the conclusion that receiving external backing affects discourse, and not the other way around.

The moderation analysis uses a binning estimator in a fractional logistic model (Hainmueller, Mummolo, and Xu 2018), with the moderator of monthly global oil price and the multiplicative interaction term of oil price and the origin of support. The model compares (1) how the conditional marginal effect of support from SA or other Arab Gulf states on radical rhetoric changes across oil prices to (2) how the effect of support from other patrons changes across oil prices. The observations of oil price are monthly, so we do not use the annual measures of group casualties and number of fighters as covariates. Instead, we use group fixed effects, which our initial logistic models show produce coefficients in the same direction and at the same level of significance as when using the covariates (see Appendix F).³⁴ For further methodological details of our analytic strategy, see Appendix D.

3.4 Results

We find a positive and significant relationship between receiving external financial support and groups being more likely to espouse the rhetoric of subversion. Table 3 presents mean marginal effects, with no support as the base level. The bivariate model indicates that receiving funding is associated with an increase of .14 ($p < .01$) in our rhetoric ratio. With the mean rhetoric ratio about .66, this increase is a substantial shift toward a radicalism of subversion. Including a group’s casualties and number of fighters as covariates does not change the coefficient for external support. We also find that experiencing casualties slightly decreases the probability of espousing a radicalism of subversion (.02, $p < 0.05$), and a group’s membership strength has no significant effect. The findings regarding external support are robust to the inclusion of fixed effects and the removal of prominent groups from the dataset (Appendix F).

³²As Högbladh et al. (2011) note, “The resistance groups in Afghanistan were supported by many governments and private donors; although many of them did not openly admit that they were giving aid to the *mujahideen* ... Saudi Arabia occasionally flew in arms deliveries to airports in Rawalpindi and Islamabad ... There are also reports that the [Arab] Gulf states were providing the guerrillas with financial aid. However, large parts of the funds seem to have been donated by individuals [residing in SA and Gulf states].” See also Roy 1986 and Rubin 2002.

³³In our primary model, we do not lag the price of oil. We assume actors in SA and the Arab Gulf states decide to offer support based on their expectations of wealth derived from current oil prices. However, because we do not have strong evidence in support of this assumption, we also estimate our model using oil prices lagged by one month and six months. The overall finding remains the same (Appendix F).

³⁴In a subsequent robustness check, we include data on groups’ geographical areas of militant activity. The results are similar

Table 3: Effect of External Support on Rhetoric of Subversion

	(1) Rhetoric Ratio	(2) Rhetoric Ratio
External support	0.138** (0.0445)	0.141** (0.0384)
Group casualties (100s)		-0.0194* (0.00860)
Number of fighters (1,000s)		0.000189 (0.000445)
Observations	12,802	3,176
Pseudo R^2	0.007	0.017

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Note: Coefficients are mean marginal effects derived from a fractional logit model. Standard errors, in parentheses, are clustered by group and period. Results are based on two-tailed tests.

These results provide evidence in support of our hypothesis, although they could be due to the rhetoric attracting more external support. To adjudicate the direction of the association, we estimate the effect of receiving support from SA or the Arab states of the Persian Gulf region on groups' rhetoric, compared to the effect on groups without this support, moderated by global oil price.

The results from the moderation analysis show that receiving external support leads to the adoption of rhetorical type. When the oil price is low—when it is least likely to result in a differential effect between support from SA or Arab Gulf states and other patrons—we find no statistically significant differences in the effect of support from Gulf countries on rhetorical type relative to non-Gulf countries (Figure 3). However, as the oil price increases, the conditional mean marginal effect of external support from SA or Arab Gulf states on the rhetoric of subversion becomes positive and statistically significant.³⁵ This change across the values of the exogenous moderator supports the hypothesized causal direction. We report the coefficients of the constitutive and interaction terms in Appendix E.

Figure 3 also allows us to examine the validity of the modeling assumptions (Hainmueller et al. 2018). Namely, we find a linear effect across values of the moderator, and the conditional effect estimates of the binning estimator reflect the standard multiplicative interaction model results. In addition, there is common support for the independent variable and moderator across levels of oil price, particularly at the mid-range values. These diagnostics strengthen our confidence in the findings.

to those in the primary analysis (Appendix F).

³⁵Results presented in Appendix F show that the findings are robust to groups' geographical area of activity and removing observations that were originally included based on the assumption that Arab fighters in a group represent ties to the Arab states of the Persian Gulf. Appendix F also presents similar results when lagging the oil price by one month and six months.

Figure 3: Effect of Saudi or Gulf Support on Radical Rhetoric of Subversion, Moderated by Price of Oil

Note: The dark line and gray 95 percent confidence interval band depicts the conditional marginal effect of support from SA or Arab Gulf states across oil price levels estimated by the standard multiplicative interaction model. The point estimates with 95 percent confidence interval bars represent the conditional marginal effects from a binning estimator at the medians of the three moderator-value bins, the first tercile (L), the second tercile (M), and the third tercile (H). The stacked histogram shows the distribution of the moderator (oil price in US Dollars; 2005 = \$100), with the red and white shaded bars indicating the distribution of the moderator in the treatment and control groups, respectively.

4 Discussion

To advance our understanding of radical discourse, we identified and analyzed radicals’ rhetoric across two decades in Afghanistan. We found that radical discourse emerges from an interplay between rhetorics, such as arguments for changing society through attacks against perceived outsiders (*i.e.*, a radical rhetoric of subversion) or through an inwardly focused remaking of individuals and communities (*i.e.*, a radical rhetoric of reversion). In addition, the balance between rhetorics reflects radicals’ social relationships. Radicals’ avenues of support will generate encounters between their own social domain and the domain of their backers (White 2008). The resulting inter-domain space—what Mische and White (1998) term a “public”—engenders an interaction that constrains and guides the construction of ideas and arguments, or rhetoric. Radical discourse thus develops through relational processes comprising interactions between radicals and supporters over time, mirroring the relational mechanisms driving radical behavior and tactics (della Porta 2013; Alimi

et al. 2015; della Porta 2018).

4.1 Contributions and future research

Our conceptual work represents an initial step toward theorizing radical discourse, but we are encouraged by the empirical evidence of rhetorical varieties and the dynamics behind their adoption. Future research could draw on our insights to develop discursive explanations for radicalization and radical action. That is, which varieties of rhetoric, or mixtures of rhetorics, appeal to which kinds of individuals? And, how do different types of rhetoric affect these individuals' behavior?

The research on radicalization increasingly suggests that the search for belonging motivates many individuals to join radical groups (Roy 2004; Bosi and della Porta 2012; della Porta 2018; Mitts 2019). Often, a sense of belonging is achieved through social relationships (Sageman 2004; Wiktorowicz 2005) and ideological resonance (Hegghammer 2010; Borum 2011; Gutiérrez Sanín and Wood 2014; Walter 2017b; Mitts 2019). Yet, it is rhetoric that often helps add meaning to relationships and shapes an individuals' understanding of a radical group's ideology which itself frequently morphs in response to strategic concerns (Goldstone et al. 1991; Christia 2012; Walter 2017a).

Thus, a focus on rhetoric could shed light on how radicals use language to construct particular experiences of belonging, helping to explain why certain communities, and not others, provide the home that radicalizing individuals are searching for. Moreover, it opens the door for exchanges across disciplinary subfields. For example, recent findings in cultural sociology (*e.g.*, Lizardo 2006; Lewis, Gonzalez, and Kaufman 2012; Lewis and Kaufman 2018) can inform research on how rhetorics of subversion and reversion shape the formation of social ties. Such work would help lay the foundation for a new research agenda on the role of culture in violent conflict and contentious politics.

Future research can also examine the unintended consequences of supporting radicals. As radicals craft their rhetoric in the public shared with patrons, we expect they will typically converge with their backers' values and perspectives (Snow et al. 1986; Snow and Benford 1988). However, this might not always be the case. Radicals, like all individuals, enter publics with their own social domains in tow, and these domains might inadvertently spur tensions between radicals and their supporters. Some radicals may in fact repudiate supporters, whom they might see as beneficiaries of the social systems they seek to change (Ferree 2003).

Our empirical analysis suggests how this potential paradox can unfold. The groups we studied that received external support were also those that advocated for a history-spanning, civilizational fight potentially threatening to their patrons (*i.e.*, a radicalism of subversion). Indeed, some of the *mujahideen* groups that the United States supported in the 1980s—groups that expectedly adopted a rhetoric championing an upending of the world order—challenged their benefactors during the first Persian Gulf War (Desert Storm) by expressing support for Saddam Hussein.³⁶ We find further suggestive evidence of this paradox when we analyze how

³⁶ *Afghanistan Information Center Bulletin*, February and August 1991.

groups' adoption of rhetoric affects future external support: receiving external support leads to adoption of the subversion rhetoric, but an increase in subversion rhetoric is not significantly associated ($p > .05$) with receiving external support during the following year (Appendix G).

Examining how networks of support can shape radical discourse in unexpected ways also has the potential to help explain the puzzling proliferation of transnational violent jihadism in recent decades (Walter 2017b). Namely, as petrostates in the Arab world grew wealthier, their citizens and governments increased support for groups espousing doctrines of personal and familial morality (*i.e.*, rhetoric of reversion) they found agreeable. But, as our findings suggest, this support may have inadvertently helped transform some of these groups into transnational Islamic militants.

Finally, future research could build on our conceptualization of radical rhetoric to develop a new theory of radicalism, as well as a broader framework for understanding the similarities and differences between various discourses of contention. To reprise, the rhetorics of subversion and reversion contrast by how they fall along dimensions of *connectivity* and *reach* (Mische 2009). That is, they are characterized by how they urge radicals to achieve their goals (connectivity), as well as by the timescale offered for achieving the desired societal transformation (reach).

The dimensions draw our attention toward a theoretical understanding of radical discourse. Rhetoric that answers questions of connectivity and reach constructs a projection of the future (Mische 2009; 2014). In the case of radical discourse, the projected future excludes specific elements of present-day society: subversion rhetoric erases disdained powerful individuals, organizations, or institutions, and reversion rhetoric erases individuals and groups that do not undertake the advocated inward transformation. So, while many discourses offer visions of the future, such as the utopian thinkers of industrializing Europe (Calhoun 2012), we posit that radical discourse is typified by its clear articulation of a boundary of exclusion from the future. Radical discourse makes explicit that some segment of society—or its ideas or behavior—is unwelcome in an imagined future.

Moving beyond radical discourse, considering its rhetorical dimensions also helps us distinguish it from similar discourses, such as revolutionism. Whereas revolutionism primarily focuses on mass contention and transformation of sociopolitical institutions (Beck 2015), radicalism can interweave an intimate personalized and localized struggle into fundamental social change. Furthermore, the notion of rhetorical boundaries suggests a boundary perspective for interpreting discourses of contention as a whole. After all, a large body of research on these discourses note their exclusionary nature (Rodgers 1987; McAdam, Tarrow, and Tilly 2001; Tarrow 2013). For example, nationalist discourse describes the political boundaries people believe they deserve by identifying who is permitted to participate within the polity (Hechter 2000; Bonikowski and DiMaggio 2016). Populism defines the bounds of the “common people” by excluding a corrupt elite, detestable lower classes, and other outsiders (Mudde 2004; Jansen 2011; Bonikowski and Gidron 2016; Müller 2016; Brubaker 2017). Fascist speech excludes the unvirtuous from the imagined nation (Stanley 2018), which, intriguingly, often uses a rhetoric of the *past* to depict the nation as having been always

virtuous.

Developing a theory of contentious discourse based on the metaphor of boundaries could help illuminate how these discourses are distinct, as well as what they share in common. In addition, it can help us analyze how discourses merge. For example, it might become apparent that “radical nationalist” discourse entails the invocation of both a political boundary and a boundary of futurity. Radical nationalist discourse would thus be understood as a discourse excluding people from the contemporary politics of a nation, as well as from the nation’s future.

4.2 Limitations

Our study is not without limitations. We address one of these, potential biases due to selective inclusion of historical data, by building the most expansive corpus possible and assessing its coverage. Moreover, this limitation—common when using historical data—is a trade-off with the opportunity to mitigate common response errors. Historical data like ours are not biased by social desirability or faulty memory; we can accurately capture the rhetoric that radical groups chose to espouse.

A second limitation is also related to the data type: achieving causal identification when using observational data. With this in mind, we examine the direction of the association between support and rhetoric. Using an exogenous variable, global oil price, in a moderation analysis, we offer evidence that external support affects which variety of radical discourse is adopted, and not the other way around. These findings, however, do not rule out unobserved confounders that could be responsible for our main result. But, such unobserved variables would have to be correlated with external support and the rhetoric ratio, and have a similar relationship with external support as a function of oil price. Because this condition is relatively difficult to satisfy, we see a moderation analysis as a satisfactory way to understand the connection between support and rhetorics without the strong assumptions necessary in alternative approaches.

Finally, a third limitation, the scope of our insights into rhetoric, needs to be addressed with further research. We do not examine the narrower scope of how case-specific cultural symbols and references are used in the rhetoric. This would require an immersion into Deobandi and Salafist doctrine, as well as the writings of thinkers and ideologues like Khomeini, Mawdudi, Qutb, and Faraj. Instead, we purposefully maintain a relatively idealized perspective to be able to identify and conceptualize general forms of radical rhetoric.

We also do not examine the broader scope of organizational rhetoric. We aimed to analyze and interpret variation within radical discourse, not how radical groups use language in different ways than, say, civil society groups working within established norms and institutions. A comparison of rhetoric across different kinds of social organizations would require expanding our corpus to include texts from non-radical organizations.

5 Conclusion

What kinds of rhetoric pattern radical discourse, and why do radicals adopt one kind over another? To answer these questions, we conducted a computational abductive analysis of an original corpus of more than 23,000 pages published by Afghan radical groups between 1979 and 2001.

We find two kinds of rhetoric within radical discourse. One, a rhetoric of subversion, targets enemies very distinct from radicals and situates these enemies in an ongoing historically-scaled and inter-communal struggle. The second, a rhetoric of reversion, is intimate. It turns inward, uprooting prevailing ideas and ways of life through the judging and policing of personal and local morality and behavior. These findings draw attention to a commonly overlooked aspect of radicalism: its everyday, ahistorical, and personal and local gaze. In other words, radicalism is often conceived as an outward-facing attack against powerful, well-established external ideas and actors (*e.g.*, Ferree 2003; Calhoun 2012; Beck 2015), but we identify rhetoric that urges radicals to profoundly change society by turning toward their immediate social and temporal context—including themselves—in an effort to exclude unwanted beliefs, behavior, people, and organizations.

Our findings also illuminate how radical rhetoric falls along dimensions of connectivity and reach (Mische 2009; 2014). Through the former, rhetoric depicts a way for radicals to reach their goals—by offensive attacks or socially intimate transformation—and, through the latter, it defines a timescale—ranging from the epochal to a focus on the immediate past, present-day, and near future. Moreover, by articulating connectivity and reach, rhetorics of radicalism construct a projection of the future that excludes specific ideas, people, groups, and institutions present in a current society. Building on this insight, we suggest that radical discourse might be usefully theorized as a discourse that rhetorically crafts an exclusionary boundary of futurity.

Next, we arrive at an answer for our second question. A radical rhetoric of subversion emerges when radicals receive support from extra-local sources and, as a result, have weaker connections with their surrounding local communities. In contrast, radicals are more likely to espouse a rhetoric of reversion when groups have less external patronage and, consequently, a stronger relationship to local communities. Our analysis shows how social relationships and social domains can impose constraints on the development and adoption of rhetoric, but also allow for unexpected tensions between radicals and their supporters.

Alongside contributions to the computational analysis of discourse, our study offers a complement to the relational perspective on radicalization (della Porta 2013; Alimi et al. 2015; della Porta 2018). Namely, our findings point toward an avenue of research on how the dynamics of social relationships and discursive boundaries shape the language of radicalism. Extending this relational approach to related discourses—populism, nationalism, and revolutionism, for example—can help us better understand how discourses of contention construct different kinds of boundaries but, at the same time, share potential to engender new forms of exclusion.

References

- Achilov, Dilshod and Sedat Sen. 2017. "Got Political Islam? Are Politically Moderate Muslims Really Different from Radicals?" *International Political Science Review* 38:608–624.
- Alimi, Eitan, Chares Demetriou, and Lorenzo Bosi. 2015. *The Dynamics of Radicalization*. New York: Oxford University Press.
- Anderson, Elisabeth. 2018. "Policy Entrepreneurs and the Origins of the Regulatory Welfare State: Child Labor Reform in Nineteenth-Century Europe." *American Sociological Review* 83:173–211.
- Bartlett, Jamie and Carl Miller. 2012. "The Edge of Violence: Towards Telling the Difference Between Violent and Non-Violent Radicalization." *Terrorism and Political Violence* 24:1–21.
- Beardsley, Kyle, Kristian Skrede Gleditsch, and Nigel Lo. 2015. "Roving Bandits? The Geographical Evolution of African Armed Conflicts." *International Studies Quarterly* 59:503–516.
- Bearman, Peter. 1993. *Relations into Rhetorics: Local Elite State Structure in Norfolk, England, 1540-1640*. New Brunswick, NJ: Rutgers University Press.
- Beck, Colin. 2015. *Radicals, Revolutionaries, and Terrorists*. New York: Polity.
- Benjamin, Daniel and Steven Simon. 2002. *The Age of Sacred Terror*. New York: Random House.
- Blau, Peter. [1964] 2008. *Exchange and Power in Social Life*. New Brunswick: Transaction Publishers.
- Blei, David, Andrew Ng, and Michael Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3:993–1022.
- Bonikowski, Bart and Paul DiMaggio. 2016. "Varieties of American Popular Nationalism." *American Sociological Review* 81:949–980.
- Bonikowski, Bart and Noam Gidron. 2016. "The Populist Style in American Politics: Presidential Campaign Discourse, 1952–1996." *Social Forces* 94:1593–1621.
- Borum, Randy. 2011. "Radicalization into Violent Extremism I: A Review of Social Science Theories." *Journal of Strategic Security* 4:7–36.
- Bosi, Lorenzo and Donatella della Porta. 2012. "Micro-Mobilization into Armed Groups: Ideological, Instrumental and Solidaristic Paths." *Qualitative Sociology* 35:361–383.
- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14:63–82.
- Breiger, Ronald, Robin Wagner-Pacifici, and John Mohr. 2018. "Capturing Distinctions while Mining Text Data: Toward Low-Tech Formalization for Text Analysis." *Poetics* 68:104–119.
- Brubaker, Rogers. 2015. "Religious Dimensions of Political Conflict and Violence." *Sociological Theory* 33:1–19.
- Brubaker, Rogers. 2017. "Why Populism?" *Theory and Society* 46:357–385.
- Calhoun, Craig. 2012. *The Roots of Radicalism: Tradition, the Public Sphere, and Early Nineteenth-Century Social Movements*. Chicago: University of Chicago Press.
- Callimachi, Rukmini. 2016. "For Women Under ISIS, A Tyranny of Dress Code and Punishment." *The New York Times* Retrieved February 4, 2019:<https://www.nytimes.com/2016/12/12/world/middleeast/islamic-state-mosul-women-dress-code-morality.html>.

- Callimachi, Rukmini. 2017. "After ISIS, Smoking Openly to Feel Free." *The New York Times* Retrieved February 4, 2019:<https://www.nytimes.com/2017/04/22/world/middleeast/after-isis-smoking-openly-to-feel-free.html>.
- Christia, Fotini. 2012. *Alliance Formation in Civil Wars*. New York: Cambridge University Press.
- Cook, Karen, Coye Cheshire, and Alexandra Gerbasi. 2006. *Contemporary Social Psychology Theories*, chapter Power, Dependence, and Social Exchange, pp. 194–216. Stanford: Stanford University Press.
- Dalgaard-Nielsen, Anja. 2010. "Violent Radicalization in Europe: What We Know and What We Do Not Know." *Studies in Conflict & Terrorism* 33:797–814.
- della Porta, Donatella. 2013. *Clandestine Political Violence*. New York: Cambridge University Press.
- della Porta, Donatella. 2018. "Radicalization: A Relational Perspective." *Annual Review of Political Science* 21:461–474.
- Denny, Matthew and Arthur Spirling. 2018. "Text Preprocessing for Unsupervised Learning: Why It Matters, When It Misleads, And What to Do About It." *Political Analysis* 26:168–189.
- DiMaggio, Paul, Manish Nag, and David Blei. 2013. "Exploiting Affinities Between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding." *Poetics* 41:570–606.
- Emirbayer, Mustafa. 1997. "Manifesto for a Relational Sociology." *American Journal of Sociology* 103:281–317.
- Ferree, Myra Marx. 2003. "Resonance and Radicalism: Feminist Framing in the Abortion Debates of the United States." *American Journal of Sociology* 109:304–344.
- Fortana, Virginia Page, Nicholas Lotito, and Michael Rubin. 2018. "Don't Bite the Hand that Feeds: Rebel Funding Sources and the Use of Terrorism in Civil Wars." *International Studies Quarterly* 62:782–794.
- Friedland, Roger. 2002. "Money, Sex, and God: The Erotic Logic of Religious Nationalism." *Sociological Theory* 30:381–425.
- Gambetta, Diego and Steffen Hertog. 2016. *Engineers of Jihad*. Princeton, NJ: Princeton University Press.
- Goldman, Marion and Steven Pfaff. 2018. *The Spiritual Virtuoso*. New York: Bloomsbury.
- Goldstone, Jack, Ted Gurr, and Farrokh Moshiri. 1991. *Revolutions of The Late Twentieth Century*. New York: Westview Press.
- Gould, Roger. 1995. *Insurgent Identities*. Chicago: University of Chicago Press.
- Grimmer, Justin and Brandon Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21:267–297.
- Gutiérrez Sanín, Francisco and Elisabeth Wood. 2014. "Ideology in Civil War: Instrumental Adoption and Beyond." *Journal of Peace Research* 51:213–226.
- Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2018. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* DOI:10.1017/pan.2018.46.
- Hechter, Michael. 2000. *Containing Nationalism*. New York: Oxford University Press.
- Hegghammer, Thomas. 2009. *Global Salafism: Islam's New Religious Movement*, chapter Jihadi-Salafis Or Revolutionaries?, pp. 244–266. New York: Columbia University Press.
- Hegghammer, Thomas. 2010. "The Rise of Muslim Foreign Fighters: Islam and the Globalization of Jihad." *International Security* 35:53–94.

- Höglbladh, Stina, Therése Pettersson, and Lotta Themnér. 2011. "External Support in Armed Conflict 1975-2009, Presenting New Data." Technical Report Montreal, Canada, 52nd Annual International Studies Association Convention.
- Isaacs, Matthew. 2016. "Sacred Violence or Strategic Faith? Disentangling the Relationship between Religion and Violence in Armed Conflict." *Journal of Peace Research* 53:211–225.
- Jansen, Robert. 2007. "Resurrection and Appropriation: Reputational Trajectories, Memory Work, and the Political Use of Historical Figures." *American Journal of Sociology* 112:953–1007.
- Jansen, Robert. 2011. "Populist Mobilization: A New Theoretical Approach to Populism." *Sociological Theory* 29:75–96.
- Jansen, Robert. 2017. *Revolutionizing Repertoires*. Chicago: The University of Chicago Press.
- Kitchin, Rob. 2014. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* 1:1–12.
- Lewis, Kevin, Marco Gonzalez, and Jason Kaufman. 2012. "Social Selection and Peer Influence in an Online Social Network." *Proceedings of the National Academy of Sciences* 109:68–72.
- Lewis, Kevin and Jason Kaufman. 2018. "The Conversion of Cultural Tastes into Social Network Ties." *American Journal of Sociology* 123:1684–1742.
- Light, Ryan and Colin Odden. 2017. "Managing the Boundaries of Taste: Culture, Valuation, and Computational Social Science." *Social Forces* 96:877–908.
- Lipset, Seymour and Earl Raab. 1978. *The Politics of Unreason*. Chicago: The University of Chicago Press, 2nd edition.
- Lizardo, Omar. 2006. "How Cultural Tastes Shape Personal Networks." *American Sociological Review* 71:778–807.
- Loadenthal, Michael. 2013. "Deconstructing 'Eco-Terrorism': Rhetoric, Framing and Statecraft as Seen Through the Insight Approach." *Critical Studies on Terrorism* 6:92–117.
- Lucas, Christopher, Richard Nielsen, Molly Roberts, Brandon Stewart, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* 23:254–277.
- Marks, Gary, Heather Mbye, and Hyung Min Kim. 2009. "Radicalism or Reformism? Socialist Parties before World War I." *American Sociological Review* 74:615–635.
- McAdam, Doug, Sidney Tarrow, and Charles Tilly. 2001. *Dynamics of Contention*. New York: Cambridge University Press.
- McFarland, Daniel, Kevin Lewis, and Amir Goldberg. 2016. "Sociology in the Era of Big Data: The Ascent of Forensic Social Science." *The American Sociologist* 47:12–35.
- Miller, Ian Matthew. 2013. "Rebellion, Crime and Violence in Qing China, 1722-1911: A Topic Modeling Approach." *Poetics* 41:626–649.
- Mische, Ann. 2003. In *Social Movements and Networks: Relational Approaches to Collective Action*, chapter Cross-Talk in Movements: Reconceiving the Culture-Network Link, pp. 258–280. New York: Oxford University Press.
- Mische, Ann. 2007. *Partisan Publics: Communication and Contention across Brazilian Youth Activist Networks*. Princeton: Princeton University Press.
- Mische, Ann. 2009. "Projects and Possibilities: Researching Futures in Action." *Sociological Forum* 24:694–704.
- Mische, Ann. 2011. *The SAGE Handbook of Social Network Analysis*, chapter Relational Sociology, Culture, and Agency, pp. 80–98. London: SAGE, 2 edition.

- Mische, Ann. 2014. "Measuring Futures in Action: Projective Grammars in the Rio+20 Debates." *Theory and Society* 43:437–464.
- Mische, Ann and Harrison White. 1998. "Between Conversation and Situation: Public Switching Dynamics across Network Domains." *Social Research* 65:695–724.
- Mitts, Tamar. 2019. "From Isolation to Radicalization: Anti-Muslim Hostility and Support for ISIS in the West." *American Political Science Review* 113:173–194.
- Mohr, John. 1998. "Measuring Meaning Structures." *Annual Review of Sociology* pp. 345–370.
- Mohr, John and Petko Bogdanov. 2013. "Topic Models: What They Are and Why They Matter." *Poetics* 41:545–569.
- Molm, Linda and Karen Cook. 1995. *Sociological Perspectives on Social Psychology*, chapter Social Exchange and Social Networks, pp. 209–235. Boston: Allyn and Bacon.
- Moskalenko, Sophia and Clark McCauley. 2009. "Measuring Political Mobilization: The Distinction Between Activism and Radicalism." *Terrorism and Political Violence* 21:239–260.
- Mudde, Cas. 2004. "The Populist Zeitgeist." *Government and Opposition* 39:541–563.
- Müller, Jan-Werner. 2016. *What is Populism?* Philadelphia: University of Pennsylvania Press.
- Nasr, Seyyed Vali Reza. 2005. "The Rise of 'Muslim Democracy'." *Journal of Democracy* 16:13–27.
- Nelson, Laura. 2017. "Computational Grounded Theory: A Methodological Framework." *Sociological Methods & Research* DOI:10.1177/0049124117729703:1–40.
- Nielsen, Richard. 2017. *Deadly Clerics*. New York: Cambridge University Press.
- Papke, Leslie and Jeffrey Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11:619–632.
- Quinn, Kevin, Burt Monroe, Michael Colaresi, Michael Crespin, and Dragomir Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54:209–228.
- Rana, Muhammad Amir. 2008. "Jihadi Print Media in Pakistan: An Overview." *Conflict and Peace Studies* 1:1–18.
- Rashid, Ahmed. 2000. *Taliban: Islam, Oil, and the New Great Game in Central Asia*. New York: I.B. Tauris.
- Rink, Anselm and Kunaal Sharma. 2018. "The Determinants of Religious Radicalization: Evidence from Kenya." *Journal of Conflict Resolution* 62:1229–1261.
- Roberts, Molly, Brandon Stewart, and Edoardo Airoidi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111:988–1003.
- Roberts, Margaret, Brandon Stewart, and Dustin Tingley. 2017. "stm: R Package for Structural Topic Models." *Journal of Statistical Software* .
- Roberts, Margaret, Brandon Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58:1064–1082.
- Rodgers, Daniel. 1987. *Contested Truths: Keywords in American Politics Since Independence*. New York: Basic Books.
- Roy, Olivier. 1986. *Islam and Resistance in Afghanistan*. New York: Cambridge University Press.

- Roy, Olivier. 2004. *Globalized Islam: The Search for a New Ummah*. New York: Columbia University Press.
- Roy, Olivier. 2017. *Jihad and Death*. New York: Oxford University Press.
- Rubin, Barnett. 2002. *The Fragmentation of Afghanistan*. New Haven, CT: Yale University Press, 2nd edition.
- Sageman, Marc. 2004. *Understanding Terror Networks*. Philadelphia: University of Pennsylvania Press.
- Salehyan, Idean, David Siroky, and Reed Wood. 2014. "External Rebel Sponsorship and Civilian Abuse: A Principal-Agent Analysis of Wartime Atrocities." *International Organization* 68:633–661.
- Saul, Heather. 2015. "ISIS 'Throw Man off a Building for Homosexual Affair' and Beat Him to Death When He Survives." *The Independent* Retrieved February 4, 2019:<https://www.independent.co.uk/news/world/middle-east/isis-throw-man-off-a-building-for-homosexual-affair-and-beat-him-to-death-when-he-survives-10012709.html>.
- Snow, David and Robert Benford. 1988. *From Structures to Action: Social Movement Participation Across Cultures*, chapter Ideology, Frame Resonance, and Participant Mobilization, pp. 197–217. Greenwich: JAI Press.
- Snow, David, Robert Benford, Holly McCammon, Lyndi Hewitt, and Scott Fitzgerald. 2014. "The Emergence, Development, and Future of the Framing Perspective: 25+ Years since 'Frame Alignment'." *Mobilization: An International Quarterly* 19:23–45.
- Snow, David and Remy Cross. 2011. "Radicalism within the Contexts of Social Movements: Processes and Types." *Journal of Strategic Security* 4:115–130.
- Snow, David, E. Burke Rochford, Jr., Steven Worden, and Robert Benford. 1986. "Frame Alignment Processes, Micromobilization, and Movement Participation." *American Sociological Review* 51:464–481.
- Snow, David, Louis Zurcher, and Sheldon Eklund-Olson. 1980. "Social Networks and Social Movements: A Microstructural Approach to Differential Recruitment." *American Sociological Review* 45:787–801.
- Somers, Margaret. 1995. "Narrating and Naturalizing Civil Society and Citizenship Theory: The Place of Political Culture and the Public Sphere." *Sociological Theory* 13:229–274.
- Somers, Margaret. 2008. *Genealogies of Citizenship*. New York: Cambridge University Press.
- Stanley, Jason. 2018. *How Fascism Works: The Politics of Us and Them*. New York: Random House.
- Stovel, Katherine and Lynette Shaw. 2012. "Brokerage." *Annual Review of Sociology* 38:139–158.
- Strick van Linschoten, Alex and Felix Kuehn. 2018. *The Taliban Reader: War, Islam, and Politics*. London: Hurst.
- Sundberg, Ralph, Kristine Eck, and Joakim Kreutz. 2012. "Introducing the UCDP Non-State Conflict Dataset." *Journal of Peace Research* 49:351–362.
- Tarrow, Sidney. 2013. *The Language of Contention*. New York: Cambridge University Press.
- Tavory, Iddo and Stephan Timmermans. 2014. *Abductive Analysis: Theorizing Qualitative Research*. Chicago: The University of Chicago Press.
- Tezcür, Güneş Murat and Clayton Besaw. 2017. "Jihadist Waves: Syria, the Islamic State, and the Changing Nature of Foreign Fighters." *Conflict Management and Peace Science* DOI: 10.1177/0738894217738990.
- Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Hunziker Philipp, and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Dataset Family." *Journal of Conflict Resolution* 59:1327–1342.

- Walter, Barbara. 2017a. "The Extremists Advantage in Civil Wars." *International Security* 42:7–39.
- Walter, Barbara. 2017b. "The New New Civil Wars." *Annual Review of Political Science* 20:469–486.
- White, Harrison. 1995. "Network Switchings and Bayesian Forks: Reconstructing the Social and Behavioral Sciences." *Social Research* 62:1035–1063.
- White, Harrison. 2008. *Identity and Control*. Princeton: Princeton University Press, 2nd edition.
- Wiktorowicz, Quintan. 2005. *Radical Islam Rising: Muslim Extremism in the West*. Lanham, MD: Rowman & Littlefield Publishers.
- Wilkerson, John and Andreu Casas. 2017. "Large-Scale Computerized Text Analysis in Political Science: Opportunities and Challenges." *Annual Review of Political Science* 20:529–544.
- Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.
- Wucherpfennig, Julian, Nils Metternich, Lars-Erik Cederman, and Kristian Skrede Gleditsch. 2012. "Ethnicity, the State, and the Duration of Civil War." *Journal of Conflict Resolution* 64:79–115.
- Young, Michael. 2002. "Confessional Protest: The Religious Birth of U.S. National Social Movements." *American Sociological Review* 67:660–688.
- Zaef, Abdul Salam. 2010. *My Life With the Taliban*. New York: Columbia University Press.

A Text data collection

Publications were collected from the Afghanistan Center Kabul University (ACKU) archive in Kabul, Afghanistan; private booksellers in Kabul; private collections in Peshawar and Islamabad, Pakistan; private collections in Mashhad, Iran; the Taliban Sources Project (Strick van Linschoten and Kuehn 2018); and the Hoover Institute archives. The publications collected from private booksellers in Kabul were acquired as original print versions. These were then scanned to produce digital copies. Those collected in Pakistan and Iran were scanned at the site of the collections. The remainder of the publications had been digitized before acquisition.

Publications were included in the corpus if they had been produced by radical individuals or groups with an operational presence in Afghanistan at any time between 1979 and 2001. Note that during this period, international observers classified Afghan groups as “traditionalist”, “moderate”, and “Islamist”. The subsequent scholarly literature largely adopted these labels (*e.g.*, Roy 1986; Rubin 2002), and we agree with many of these distinctions. For example, it is clear that some groups were seeking a new Islamist state while others were calling for the reinstatement of the exiled king (leading, *e.g.*, to labels of “traditionalist” and “moderate”).

In this study, we consider all these groups as radical because they satisfy two basic social-scientific definitions of radicalism. The first is a *goal-oriented* understanding: each group opposed hegemonic sociopolitical systems or powerful people, organizations, and institutions in a particular context (Goldstone et al. 1991; Ferree 2003; Calhoun 2012; Bartlett and Miller 2012; Beck 2015; della Porta 2018). The second is a *tactic-oriented* understanding: each group used extra-institutional means typically violent to challenge prevailing ideas or powerful actors in the pursuit of foundational social change ((Ferree 2003; Beck 2015; della Porta 2018). Because of our definition and selection of groups, our analysis only supports inferences about the rhetoric used by radicals. We cannot shed light on civil society organizations that do not meet the goal- or tactic-oriented understandings of radicalism. For example, we cannot compare the rhetorics of radicalism to the language used by charity groups that operated in the Afghan refugee camps on the Pakistani border. Doing so would require collecting material produced by such groups.

The search for publications was limited by data survival, knowledge of existence, and resources. Regarding the first limitation, it is possible that publications once existed but no longer do. This concern is mitigated by the near-complete runs of publications the corpus contains for major groups, such as *Jam’iyyat-i Islami-yi* and the Taliban. The second concern is mitigated by our search and analysis of coverage: we know which groups operated in Afghanistan between 1979 and 2001; we can search for publications by these groups in likely languages; and we can evaluate which group-years are not represented by text data in the corpus and use this evaluation to assess potential bias. The resulting corpus composition and coverage is reported in Appendix B, Table B1 and Table B2. The third limitation is, of course, common. However, a significant amount of time and financial resources were committed to data collection by the lead author over the course of three years. These included, for example, multiple trips to Afghanistan, Pakistan, and California, as well as a three-month search of private collections in Iran by a research assistant.

B Details of the Corpus

B.1 Composition of corpus

Table B1: Document Totals in Corpus by Publisher

	Publisher	Time Span	Documents (N)
1	Harakat-i Islami-yi Afghanistan	03/1984 to 11/1999	175
2	Hizb-i Islami (Gulbuddin faction)	08/1985 to 03/1997	2,925
3	Hizb-i Wahdat	05/1990 to 11/1997	632
4	Islamic Dawat	02/1981 to 07/1985	110
5	Hizbullah Afghanistan	03/1989 to 11/1989	167
6	Ittihad-i Islami	09/1983 to 01/1987	53
	Bara-yi Azadi-yi Afghanistan		
7	Jabha-yi Nijat-i Milli-yi Afghanistan	11/1988 to 01/1998	837
8	Sibghatullah Mojaddedi	02/1989 to 11/1990	1,002
9	Jam'iyat-i Islami	01/1985 to 12/2001	4,425
10	Burhanuddin Rabbani	09/1979 to 05/1996	2,317
11	Mahaz-e Milli-ye Islami	12/1986 to 04/1989	611
12	Taliban	03/1994 to 09/2001	1,890
13	Taliban Support Council	02/1995 to 07/1997	176
14	Afghan Students Abroad	04/1985 to 04/1986	363
15	Cultural Council of Afghanistan Resistance	06/1985 to 12/1992	3,128
16	Haqqani Organization	12/1990 to 01/1992	643
17	Jami'ah al Da'wah ila al-Qur'an wa al-Sunnah	01/1988 to 11/1990	1,013
18	Loqan Hakim	04/1990	181
19	Miramshah College (Haqqani Organization affiliation)	06/1991 to 08/1992	185
20	Sayed Rasool	No date	127
21	Maktab al-Khidamat	12/1984 to 05/1995	2,337
22	"Three Mujahideen Parties"	06/1982	46
			23,343 total

B.2 Text data coverage

Table B2 shows the coverage of our corpus. The table's left-hand column displays the groups known to have operated in and around Afghanistan between 1979 and 2001. Many of these groups spanned the entire period, but some, like the Cultural Council of Afghanistan Resistance and *Sazman-i-Nasr*, did not.

The list does not include *Junbish-i-Milli*, a political party and military organization originally allied with the communist regime. Although it eventually sided with the *mujahideen*, it is not considered to have espoused a radical rhetoric, as defined here, because it often switched allegiances to align itself with those in power. Moreover, its goal was widely considered to be prioritizing the interests of its leaders and members of the leaders' ethno-kinship networks, rather than deep social change. The list also does not include major alliances, such as *Shura-e Nazar*, constituted by distinct groups. In the cases of alliances, we treated each group within the alliance as separate during the computational abductive analysis. However, the list does include groups that formed out of a union of previously existing organizations, such as *Hizb-i Wahdat*. In

these cases, we analyze both the antecedent and subsequently merged groups.

Table B2 summarizes coverage by showing the number of documents per group and the number of years with data between groups establishment and 2001, the end of our case period. It is important to recall that not all groups existed until 2001; some dates of dissolution are generally known, but several are not. Table B1 contains the date ranges for the publications of the groups listed in Table B2.

The coverage of the corpus is good: it contains multiple years of data—in some cases, thousands of pages—for the majority of the known groups. Indeed, major groups like *Hizb-i Islami Gulbuddin* (n = 2,925), *Hizb-i Wahdat* (n = 632), *Jam'iyat-i Islami* (n = 6,613), *Jabha* (n = 1,720), and the Taliban (n = 2,066), as well as consequential groups like *Maktab al-Khidmat* (n = 2,337), the precursor of al-Qaeda, all have multiple years of coverage and large number of observations. The groups that do not have coverage were either comparatively small over the long term (*e.g.*, *Harakat-i Inqilab-i Islami*), merged into a larger group after a few years (*e.g.*, *Sazman-i-Nasr*, *Sepah-i-Pasdaran*), or had relatively limited territorial influence (*e.g.*, *Hizb-i Islami Khalis*). From this list, *Hizb-i Islami Khalis* perhaps had a larger influence on discourse than its territorial control would suggest, but documents produced by the group's powerful commander, Jalaluddin Haqqani (Roy 1986), are included in the corpus.

Table B2: Coverage of Known Groups (1979–2001) by the Corpus

	Standardized name	Alternative names	Date of establishment	Number of years since establishment with data coverage	Total number of documents (N)	Languages of documents
1	Cultural Council of Afghan Resistance	–	1985	6	3,128	English
2	Haqqani Organization	Jalaluddin Haqqani	1979	3	828	Arabic, Urdu
3	Harakat-i Inqilab-i Islami-yi Afghanistan	Movement of the Islamic Revolution	1979	0	0	NA
4	Harakat-i Islami-yi Afghanistan	Subh-i Danish, Islamic Movement	1979	6	175	Dari
5	Hizb-i Islami-yi Afghanistan (Gulbuddin faction)	Hezb-Hekmatyar Hezb-Gulbuddin	1979	7	2,925	Dari, English, Pashto
6	Hizb-i Islami-yi Afghanistan (Khalis faction)	Hezb-Khalis	1979	0	0	NA
7	Hizbullah Afghanistan ^a	Jibah-tu Allah	1979	1	167	Dari
8	Hizb-i Wahdat ^a	Islamic Unity Party	1989	4	632	Dari
9	Islamic Dawat ^a	–	1979	4	110	Dari
10	Ittihad-i Islami Bara-yi Azadi-yi Afghanistan	Ittihad-Sayyaf	1979	3	53	Pashto
11	Jabha-yi Nijat-i Milli-yi Afghanistan ^b	Afghan National Liberation Front	1979	10	1,720	Dari, English, Pashto
12	Jami'ah al Da'wah ila al-Quir'an wa al-Sunnah	Jamil al-Rahman	1986	3	1,103	Arabic
13	Jam'iyyat-i Islami-yi Afghanistan ^c	Jamiat-Islami, JIA	1979	16	6,613	Arabic, Dari, English, Pashto, Urdu
14	Mahaz-i Milli-yi Islami-yi Afghanistan	National Islamic Front	1979	4	611	Pashto, Urdu
15	Maktab al-Khidmat	Services Bureau, Services Office	1984	9	2,337	Arabic
16	Sazman-i-Nasr ^a	–	1979	0	0	NA
17	Sepah-i-Pasdar ^a	–	1982	0	0	NA
18	Taliban ^d	–	1994	7	2,066	Dari, English, Pashto

^aHizb-i Wahdat is formed in 1989 through the union of various predominantly Shi'ite groups, such as Hizbullah Afghanistan, Islamic Dawat, Sazman-i-Nasr, and Sepah-i Pasdaran. ^bCoverage of Jabha includes speeches by Sibghatullah Mojaddidi, its leader. ^cCoverage of Jam'iyyat includes speeches by its leader Burhanuddin Rabbani. ^dCoverage of the Taliban includes publications by the Taliban Support Council, a group based in London that published in English and Pashto.

Note: The list of all known groups and their establishment dates was created and validated using data from Roy 1986, Rubin 2002, Christia 2012, and <http://www.payam-aftab.com>.

C Methodological details of the computational abductive analysis

C.1 Preprocessing of text data

To conduct our computational textual analysis of documents in multiple languages, we prepared the documents in a series of steps (Grimmer and Stewart 2013; Lucas et al. 2015). First, all publications were digitized and converted to plain text format. Then, the non-English publications were translated into English. Many of the originally non-English publications had already been translated by human translators, such as the entire Pashto collection from the Taliban Sources Project (Strick van Linschoten and Kuehn 2018). For the non-English documents not previously translated, we hired professional translators to translate a random selection. The selection included publications by *Jabha-yi Nijat-i Milli-yi Afghanistan* (from Dari), *Jam'iyat-i Islami-yi* (from Arabic and Urdu), the Haqqani Organization (from Arabic), *Jami'ah al Da'wah ila al-Qur'an wa al-Sunnah* (from Arabic), Miramshah College (from Arabic), and *Maktab al-Khidamat* (from Arabic). The size of the selection was limited by monetary resources.

The remaining non-English documents in the corpus were translated using `translateR`³⁷ (Lucas et al. 2015).. This process induced data loss, as is usually the case. We attempt to account for the differential rates of this data loss across our documents by including a covariate for language in our structural topic models, as explained in more detail below. We recommend that future research re-translate the computer translated documents because artificial intelligence translation tools are constantly improving. For example, as this article was accepted for publication, Microsoft released version 3 of Translator Text API; we had used version 2. To support improvements to the existing automated translations, we offer access to the non-translated versions of our documents at <https://osf.io/jx756/>.<https://osf.io/jx756/>.

After translating the entire corpus to English, the text was preprocessed (Grimmer and Stewart 2013) using the standard settings in the structural topic model package, or `stm`³⁸, for R (Roberts et al. 2016). This included splitting text strings into individual terms, converting all characters to lower case, stemming the terms, and removing special characters, numbers, stop-words, and infrequent terms.

These preprocessing changes to the words in the corpus are standard, but, as Denny and Spirling (2018) point out, each one can influence the results. To assess the impact of preprocessing on our analysis, we used the `preText`³⁹ R package (Denny and Spirling 2018). This package allowed us to estimate the effect of each preprocessing step on a “preText score”; negative coefficients mean that the step reduces the “unusualness” of results for our corpus, while positive coefficients indicate that conducting the preprocessing will likely lead to more “unusual” results.

Our preText analysis suggests that the preprocessing steps were overall unlikely to produce unusual findings (Figure C1). Four of the six steps were estimated to either have no effect or to produce more normal results. We chose to include the two steps that were likely to increase unusual findings (*i.e.*, removing punctuation and numbers) because we did not want these kinds of characters to influence the findings.

Finally, we parsed each publication into pages, which served as our observational units, or documents, for the structural topic modeling (Grimmer and Stewart 2013). Below, we evaluate the use of pages as documents (subsection C.3). At this stage, a document-term matrix was created (Blei et al. 2003) using `stm`.

C.2 Topic models: Results and interpretation

We analyze the document-term matrix with structural topic models (STM), implemented via `stm` (Roberts et al. 2016). These models include document metadata as covariates: documents’ authorship, or the radical group that published the text, the year of publication, and the original language of publication. Our main results are generated by modeling a 10 topic solution (k). The selection of $k = 10$ is explained below.

The 10 topics are shown in Figure C2 as a display of topic labels and the ten words with the highest FREX scores. FREX scores are based on terms’ probability of appearing under a topic and their exclusivity to that topic (Roberts et al. 2014). From these 10 topics, we select a subset of seven topics that are evocative of sociopolitical discourse (see section 2 in the main text): topics 1, 3, 4, 5, 6, 7, and 8.

³⁷Version 1.0; <https://cran.r-project.org/web/packages/translateR/translateR.pdf>

³⁸Version 1.3.0; <https://cran.r-project.org/web/packages/stm/stm.pdf>

³⁹Version 0.6.1; <https://cran.r-project.org/web/packages/preText/preText.pdf>

Figure C1: Regression Coefficients Across Preprocessing Steps

Finding a subset of topics to be non-relevant is not surprising. Since radical groups’ writing will not solely reference sociopolitical discourse (Rana 2008), not all topics will be relevant to our study on radicalism. For example, Topic 10 appears to capture writing about the interviews that many *mujahideen* gave to the British Broadcasting Corporation (BBC). In addition, words can be clustered into a topic simply because of their lack of fit in other topics. These “residual topics” are often difficult to interpret—this is what we see with topics 2 and 9, with words such as “pace”, “hair”, “code”, and “sunday”, “catalog”, and “friday”, respectively. For further discussion on using a subset of topics, see Nielsen 2017.

The seven selected topics are grouped into two varieties of radical rhetoric, *subversion* and *reversion*, as discussed in the main text’s subsection 2.4. The relative prominence of each rhetorical variant per document is then measured with a *rhetoric ratio score*. This score is a ratio of the prevalence of each rhetorical type’s underlying topics. That is, in each document, the prevalence of the topics in the set that defines radicalism of subversion is compared to the prevalence of the topics in the set that defines radicalism of reversion. In our case, the greater the prevalence of topics in the subversion group (*i.e.*, topics 1, 3, 5, 7, and 8), the closer a document’s rhetoric ratio is to one. The more a document is made up by topics in the reversion group (*i.e.*, topics 4 and 6), the more a score approaches zero. Finally, each radical group is given a rhetoric ratio score by calculating the mean score of their documents.

C.3 Topic models: Validation

Topic models produce a measurement of a corpus. The quality of topic models output should therefore be evaluated like other measures: by its reliability and validity (Quinn et al. 2010: 216). Because any given topic model can be repeated (Nelson 2017),⁴⁰ discussion about assessing topic model results has focused on

⁴⁰While topic models can be repeated, their multi-modal nature could lead to slightly different results (Roberts et al. 2014).

Figure C2: Topic Labels and Top FREX Words

1. War	mujahideen, kill, attack, soldier, captur, enemi, oper, command, tank, area
2. Topic 2	pepe, pace, hair, crush, bottom, fashion, mesa, dam, code, matt
3. Theology of jihad (struggle)	god, jihad, quot, messeng, allah, brother, prophet, sheikh, bless, martyr
4. Local education & markets	educ, thousand, provinc, news, train, respect, ministry, school, abdul, mullah
5. Afghan jihad (battle)	mujahedeen, jihad, soviet, najib, kabul, regim, afghan, elect, presid, pakistan
6. Religious code	movement, holy, life, great, human, scholar, women, quran, answer, world
7. International politics	afghanistan, russian, countri, govern, unit, intern, russia, support, state, american
8. Social & political revolution	social, uniti, polit, communiti, revolut, parti, west, exist, natur, system
9. Topic 9	grant, sunday, catalog, prison, friday, inevit, servic, pose, pro, innov
10. Topic 10	profil, breath, select, patent, post, bbc, threat, bar, tuesday, cost

Note: Words have been stemmed.

validity. In this section, we first present validation of our topic model results and then we offer a validation of the aggregation of topics into two rhetorical types.

C.3.1 Validation of topics

The validation of topic models typically entails assessing two related dimensions: whether the output is a “good” measurement of the corpus, which includes the selection of the topic solution (k), and whether it is credible, or the degree to which the topics “make sense” (Wilkerson and Casas 2017: 533). We discuss each in turn.

There is no right answer to the number of topics that should be modeled (Roberts et al. 2014). As Grimmer and Stewart (2013: 270) note, “the complexity of language implies that all methods necessarily fail to provide an accurate account of the data-generating process used to process texts. Automated content analysis methods use insightful, but wrong, models of texts to help researchers make inferences from their data.” This is akin to there being no correct way to measure, say, “poverty”. In studies examining poverty, selecting a measure of poverty follows from the theoretical insight that can be gained by the measure, as well as that insight’s relevance to the research question.

With this property of topic models in mind, evaluating the selection of how many topics to model should be based on whether it provides “new and useful categorization schemes for text” (Grimmer and Stewart

2013: 270). Or, as Quinn et al. (2010) put it, does the solution, which offers a view of the data at a particular granularity (Roberts et al. 2014: 6), “work”? By this criterion, our selection of a topic solution (k) was good because it helped us develop a new understanding of radicalism.

To provide further confidence in the selection of a topic solution, the common practice has been to use expert knowledge to assess the output of different specifications of k (see Grimmer and Stewart 2013). For example, Quinn and colleagues (2010) and Lucas and colleagues (2015) recommend that researchers fit models with small, medium, and large values of k , and then “qualitatively” evaluate the usefulness of the varying results.

We followed this advice by making use of the groups’ ratio scores. As discussed in the main text and in Appendix A, we consider all the groups that produced the texts in our corpus as radical. However, the scholarship on Afghanistan has long noted differences in the political objectives of these groups (*e.g.*, Roy 1986; Rubin 2002; Christia 2012). For example, some supported the reinstatement of Afghanistan’s exiled king while others advocated for the creation of an Islamist state. Thus, the rhetoric ratio score should differ somewhat between some groups. For instance, “traditionalist” radicals should not have the same score as “Islamist” radicals. We observed such logical differences in the relative scores based on the $k = 10$ model specification. Using k values of 7 and 13 produced results not in alignment with our knowledge, nor scholarship, of the case: the resulting group rhetorical ratio scores indicated the groups were largely indistinguishable from one another, which clearly contradicted the historical record. Consequently, we settled on an STM using a 10-topic solution.

Furthermore, the granularity of the topics estimated via a 10-topic solution was appropriate for conceptualizing thematic patterns and types of rhetoric within the discourse of radicalism, a central goal of our study. Fewer topics (*i.e.*, $k = 7$) resulted in too broad topics—distinct concepts were “squished” together—whereas a solution greater than 10 (*i.e.*, $k = 13$) produced topics that would have been more meaningful if combined with one another—a too fine-grained measurement of the corpus. In other words, the 10-topic solution provided a view of the corpus that helped us achieve our research goal, a key criterion for evaluating a measure’s application (Quinn et al. 2010; Grimmer and Stewart 2013; Roberts et al. 2014; Nelson 2017).

Recent studies have supplemented the qualitative validation approach with a data-driven selection of k . One technique, proposed by Roberts et al. (2014) and recently implemented by Light and Odden (2017), compares the exclusivity and semantic coherence of topics for different solutions. Exclusivity refers to the extent to which frequent words are contained in a single topic. Coherence is the tendency for topics’ most-probable words to co-occur together (Roberts et al. 2017). Topic solutions that generate higher exclusivity and coherence scores are understood to produce topics that are more semantically useful—they are consistent while also different from one another (Roberts et al. 2014).

Figure C3 shows the relationship between exclusivity and coherence for various candidate models of our corpus. We see that the 10-topic solution produces some of the most semantically useful results. A five-topic solution generates more coherent but less exclusive topics. Moving from five to 10 topics results in a large increase in exclusivity with a relatively small decrease of coherence. In contrast, moving from 10 to 15 topics results in a large reduction in coherence with relatively little increase in exclusivity. Continuing to increase the value of k incrementally increases the exclusivity, but, again, by weakening coherence. In practice, this means high-frequency words would be spread across ever more topics—words that otherwise might better belong in the same topic.

Our validation procedures give us confidence in our topic solution. A qualitative assessment based on the historical record of Afghanistan indicates that a 10-topic solution is best, as does the ultimate usefulness of this measurement. Moreover, an evaluation of candidate solutions’ exclusivity and coherence—an evaluation conducted *after* we qualitatively selected our solution—also suggests the 10-topic solution efficiently generates one of the most semantically useful views of the corpus.

A second key issue when evaluating the quality of measurement is the selection of units, or documents. How might parsing the text into, say, paragraph- or page-length documents affect our inferences? The rule of thumb is to select a text length in which there is meaningful discontinuity (Grimmer and Stewart 2013). In our corpus, however, doing so is not straightforward. Our dataset contains a variety of kinds of text, throughout which there is no consistent use of break point. For example, some printed speeches run for pages without paragraph breaks. In light of this, we used page breaks to demarcate our documents, reasoning that, first, the length of a page roughly captured the ideas the authors opted to place together and, second, every publication in our corpus had page breaks.

Figure C3: Exclusivity and Semantic Coherence of Topics for Different Topic Solutions

Note: Each point shows the exclusivity and coherence for a given topic solution, or number of topics. The topic solutions are indicated by the numbers assigned to each point.

To assess how the use of page-demarcated units might influence our findings, we first compare our main results to the results from a topic model drawing on documents of 500-word segments—a length of text that is about a page long, but not marked by page layouts. Table C3 presents the top 10 FREX words generated by these two models, with the main results listed in the second column. We observe consistency in the topics. In addition, we find largely similar FREX words rankings, although the extent of change in the rankings varies across topics.

Having established that models using documents defined as a page or 500 words produce similar results, we turn to assessing the output produced from documents of other lengths. Because, as mentioned earlier, not all publications in our corpus contain paragraph breaks, we create documents that are 150 and 350 words long and use these in the model. Then we compare the exclusivity and semantic coherence of these models' results to those of existing results. Figure C4 shows that modeling a 10-topic solution with documents 500 words long, or about a page length, maximizes exclusivity and coherence. Comparing the results shown in Figure C4 to those in Figure C3 indicates that using page-demarcated documents produces topics that are more semantically coherent but slightly less exclusive than topics generated using 500- and 350-word long documents.

Finally, we evaluate the credibility of our topic model results. To do so, we conduct a test of predictive validity (Quinn et al. 2010). Results are high in predictive validity if they vary along features external to the modeling process in expected ways. For example, we gain confidence in the credibility of topics found in U.S. senators communication with constituents if the prevalence of specific topics in a senator's communications corresponds to that senator's committee assignment (Grimmer and Stewart 2013).

Our test compares the estimated prevalence of topics across radical groups. If our results are credible, the proportion of certain topics should differ based on the groups activities. For example, a topic that captures

Table C3: Comparison of Top FREX Words

Topics	Page-length documents	500-word documents
War	mujahideen, kill, attack, soldier, capture, enemy, command, tank, area	mujahideen, kill, soldier, attack, capture, enemy, tank, operation, destroy, wound
Theology of jihad (struggle)	god, jihad, quote, messenger, alah, brother, sheikh, bless, martyr	allah, god, prophet, messenger, brother, jihad, almighty, bless, heart, martyr
Local education & markets	education, thousand, province, news, train, respect, ministry, school, abdul, mullah	education, train, children, studies, school, women, university, student, institution, province
Afghan jihad (battle)	mujahideen, jihad, soviet, najid, regime, afghan, election, president, pakistan	mujahideen, soviet, jihad, afghan, regime, kabul, refuge, najib, peshawar, volunteer
Religious code	movement, holy, life, great, human, scholar, woman, quran, answer, world	human, social, societies, communities, right, system, principle, mean, nature, life
International politics	afghanistan, russian, countries, govern, unit, international, russia, support, state, american	russian, movement, war, russia, muslim, defeat, struggle, face, armies, victory
Social & political revolution	social, unity, political, communities, revolution, parties, west, exit, nature, system	parties, council, elect, leader, iran, unit, leadership, solution, member, decision

Note: Table shows the top 10 FREX words generated from two models using different document definitions. The second column presents the main results: topics and words generated from a model drawing on page-demarkated documents. The third column presents the results from a model drawing on documents 500 words in length. Endings have been added to stemmed terms based on the usage of words in highly associated documents.

battlefield violence should be more prevalent in the writing by a militant group routinely engaging in armed fighting.

We test three topics across three groups, the Cultural Council of Afghanistan Resistance (CCAR), *Jam'iyat-i Islami*, and *Hizb-i Islami* (Gulbuddin). We selected these groups because our corpus contains similar amounts of their publications across time. The first topic, War, which comprises language about battlefield violence (*e.g.*, “kill”, “attack”, “soldier”, “tank”), should be most prevalent in the documents produced by *Jam'iyat*, the group that was the most active in the fighting (Rubin 2002). The second topic, “Theology of jihad”, made up of high-frequency words like “god”, “message”, and “prophet”, should be more common in the writings of *Hizb-i Islami*. Among our selected groups, *Hizb-i Islami* most prioritized its Islamist agenda. As Roy (1986: 133) observed at the time, its “power base did not extend beyond the network of Islamists” and “it regarded the Islamic revolution as being more important than the war” against the communists. Finally, the third topic, “Local education & markets”, should be most prevalent in the CCAR’s publications. The CCAR was not militarized; it focused on political and social issues.

Figure C5 presents the results of our test. Our topic model output varies in the ways we would expect. The “War” topic is more often estimated to have a greater prevalence in the writings of *Jam'iyat*; the

Figure C4: Exclusivity and Semantic Coherence of Topics for Different Document Lengths and Topic Solutions

Note: Each point shows the exclusivity and coherence for a given topic solution, or number of topics. The topic solutions are indicated by the numbers assigned to each point. Colors indicate the relationship for topic solutions when using documents of 500, 350, and 150 words long.

estimated proportion of “Theology of jihad” is most frequently greater in the documents produced by *Hizb-i Islami*; and the “Local education & markets” estimate is most often greatest in CCAR’s publications. In summary, the assessments of measurement quality and credibility together give us confidence in the validity of our topic model results.

In sum, the assessments of measurement quality and credibility together give us confidence in the validity of our topic model results.

C.3.2 Validation of rhetorical variants

We also validate the aggregation of our topics into two rhetorical varieties using the criteria of credibility (*i.e.*, does it “make sense”?) and quality of measurement (*i.e.*, does it capture an important aspect of the data well?) (Wilkerson and Casas 2017: 533). The assessment of the latter is provided by the theoretically driven computational abductive analysis, as explained in the main text. This section is used to evaluate the former.

As in the preceding section, we check the aggregation’s credibility using predictive reasoning: the results of the aggregation should correspond to external events, categories, or findings in expected ways (DiMaggio, Nag, and Blei 2013; Grimmer and Stewart 2013). Because one hallmark of our rhetorical types is that they differ in how they portray radicals’ *connectivity* to the future, or the logic by which radicals move toward their imagined future (subsection 2.4 in the main text), we compare their variation in language to the variation observed in other depictions of connectivity.

In her study of language used by environmental policy makers and activists, Mische (2014) identified sets of words associated with kinds of connectivity. One set, “advance”, “against”, “attack”, “fight”, and “must”, indicates a call to action—an engagement in a struggle to reach a certain future. This kind of connectivity is akin to our radicalism of subversion. A second set of words suggests that a future will happen under certain conditions, similar to how a radicalism of reversion encourages listeners to turn inwards towards themselves and their families and communities to engender a desired future. These words are, “can”, “might”, “likely”, “might”, “will”, and “would”. We use a selection from the first and second sets of words to denote rhetorics of subversion and reversion, respectively.

Our validation test entails two kinds of comparisons. First, we compare the frequency of the exogenously defined subversion terms in the subversion text to their frequency in the reversion text. At the same time, we compare the frequency of the reversion terms in the subversion text to their frequency in the reversion text. If we have developed a good measurement of our theorized rhetorics, the subversion terms should be more frequent in the subversion text, and the reversion terms should be more frequent in the reversion text. In addition, we conduct these word frequency comparisons with random aggregations of topics, or what can be thought of as placebo rhetorics.

Second, we compare (1) how well the word frequency comparisons with our theoretically justified aggregations align with expectations to (2) how well the word frequency comparisons done with placebo rhetorics align to expectations. The comparisons done with the theoretically justified aggregations—the rhetorics of subversion and reversion—should fit expectations better than the comparisons using the placebo aggregations.

To conduct our test, we first identify the documents that have an estimated topic proportion greater than 75% for each topic. These documents are then merged into a single corpus, which is subsequently preprocessed as described in subsection C.1 and subset to only contain common English words. The overall results remain the same if the word frequencies are normalized using the entire corpus. We then combine the documents highly associated with topics 1, 3, 5, 7, and 8 into a subversion text. The documents highly associated with topics 4 and 6 are compiled into a reversion text.

The placebo texts are formed by compiling all possible random combinations of topics. For example, the documents highly associated with topics 1, 5, and 6 might be combined into a “rhetoric”, captured by a “T1-T5-T6” text. The rhetorics reflected in the placebo texts might be arbitrarily labeled as “subversion” for ease of interpretation, but they could have been labeled as examples of reversion. We present the results of example placebos in Table C4, but we also calculate the mean result for all placebos.

Table C4 shows the results of our comparisons. The word frequencies are percentages of the total words in each aggregation’s compiled text. For example, a value of .01 for the term “advance” in the subversion compilation of text indicates that .01 of all words in the compilation are “advance.” The frequencies of our selected terms differ across the theoretically justified subversion and reversion rhetorics in the expected ways. All the subversion terms are more common in the subversion rhetoric than in the reversion rhetoric. Similarly, all the reversion terms appear more frequently in the reversion text than in the subversion text. Some terms appear twice as often in the appropriate rhetoric.

In the second step of our test, we find that the theoretically justified aggregations differ in expected ways more often than the random aggregations. The first example placebo comparison (*i.e.*, topics 1, 5, and 6 versus topics 3, 4, 7, and 8; “Random selection 1” in Table C4) achieves three out of ten comparisons correctly, or, when the arbitrarily assigned labels are switched, seven out of ten comparisons align with expectations. In the second example placebo comparison (*i.e.*, Topics 3 and 6 versus Topics 1, 4, 5, 7, and 8; “Random selection 2” in Table C4) three out of ten expectations again with expectations, or, when the labels are switched, seven out of ten match expectations. Thus, two of the example placebos come close (70 percent success) while the other two perform poorly (30 percent success) while the rhetorics of subversion and reversion achieve 100 percent success in this test. The mean success rate of all possible aggregations is six out of 10 (60 percent) with a standard deviation of 1.41. The success rate of the theoretically justified categories is greater than one standard deviation of the placebos’ mean success rate. These results suggest that the rhetorics of subversion and reversion better align with previously observed rhetorical treatments of connectivity (Mische 2014) than other possible aggregations of our topics.

In addition, we check how using individual topics instead of aggregated rhetorics would affect our insights regarding the adoption of rhetoric (*i.e.*, the second part of our study). We find that, overall, we arrive at the same conclusion. These results are presented in subsection F.1.

Figure C5: Prevalence of Selected Topics by Group

Note: The estimated topic proportion of specific topics vary across groups in expected ways. The boxes denote the median and 25% to 75% quartile range. Points are outliers.

Table C4: Percent of Test Terms in Each Rhetorical Category

		Rhetorics		Random selection 1		Random selection 2	
		Subversion	Reversion	“Subversion”	“Reversion”	“Subversion”	“Reversion”
		Topics 1, 3, 5, 7, 8	Topics 4, 6	Topics 1, 5, 6	Topics 3, 4, 7, 8	Topics 3, 6	Topics 1, 4, 5, 7, 8
Subversion test	advance	0.01	0.01	0.00	0.01	0.00	0.01
	against	0.19	0.16	0.36	0.05	0.07	0.31
	attack	0.40	0.01	0.76	0.01	0.01	0.69
	fight	0.05	0.02	0.04	0.06	0.07	0.03
	must	0.07	0.05	0.03	0.09	0.10	0.03
Reversion test	can	0.10	0.37	0.14	0.13	0.18	0.10
	likely	0.00	0.01	0.00	0.01	0.01	0.00
	might	0.01	0.03	0.02	0.01	0.02	0.01
	will	0.32	1.01	0.48	0.38	0.43	0.42
	would	0.13	0.22	0.22	0.08	0.08	0.20

Note: The theoretically justified rhetorical categories (*i.e.*, aggregations of topics), rhetorics of subversion and reversion, have a greater number of comparative word frequencies that differ as expected than the aggregations of randomly selected topics. The values in the cells indicate the percent of all words in the rhetorical categories that are the test terms. Values have been rounded. The bold pairs of percentages across each pair of theorized or placebo rhetorics indicate a difference in the expected direction, although the placebo results should also be evaluated with labels switched. See subsection C.3 for details.

C.4 Topic models: Robustness

An important consideration when interpreting STM output is how robust the topics are to the omission of covariates. When producing our initial STM results, we did not consider how the type of publication (*e.g.*, magazine, newspaper, speech) may influence the topical content. To check for this potential effect, we re-estimated our model with the addition of publication type to the original set of covariates. We found no major differences in the output of the original and updated model. The rank ordering of FREX words in the updated model was slightly different (Figure C6), but the content of the topics remained nearly the same (compare to Figure 1 in subsection 2.3 in the main text). In light of these findings, we used the original output for the main analysis.

Figure C6: Labels and Top FREX Words for Selected Topics Estimated from STM with Publication Covariate

Topic 1. War	mujahideen, kill, solider, attack, capture, enemy, operation, command, tank, destroy
Topic 3. Theology of jihad (struggle)	god, jihad, quote, messenger, brother, allah, prophet, bless, sheikh, martyr
Topic 4. Local education & Markets	education, thousand, province, train, news, respect, abdul, kandahar, ministry, school
Topic 5. Afghan jihad (battle)	mujahideen, jihad, soviet, najib, kabul, regime, afghan, pakistan, rabbani, president
Topic 6. Religious code	movement, holy, life, great, human, scholar, women, quran, answer, live
Topic 7. International politics	russian, countries, afghanistan, unit, govern, intern, russia, state, support, american
Topic 8. Social & political revolution	social, unity, political, communities, revolution, parties, west, exist, nature, system

Note: Figure shows the content of topics estimated with an STM using the original covariates and the addition of a covariate for publication type. The content is nearly identical to the original output (Figure 1 in subsection 2.3 in the main text). Endings have been added to stemmed terms based on the usage of words in each topic's highly associated documents.

D Methodological details of the fractional logistic analysis

The initial fractional logistic models utilize two main sources of data. One, our dependent variable, is a document-level ratio of subversion to reversion rhetorics. The ratio is determined for each document in our corpus by comparing the proportion of a document’s text associated with topics underlying each rhetorical variant (see subsection C.2). In our case, the higher the value of the ratio, the more a document’s discourse reflects a radicalism of subversion. To gain further confidence in our conclusions, we subsequently redo the analysis using the prevalence of individual topics, rather than aggregated rhetorics, as the dependent variable. These results are presented in Appendix F.

The second source of data, the Uppsala Conflict Data Program (UCDP),⁴¹ informs our independent variables. Specifically, we use the UCDP datasets containing information about (Högbladh et al. 2011) and groups’ casualties and number of fighters (Sundberg et al. 2012), as well as the affiliated ACD2EPR dataset⁴² (Wucherpfennig et al. 2012; Vogt et al. 2015) for ethnic and sectarian affiliation.

Observations are annual-level because the UCDP data on our independent variable of interest—whether a group receives external financial support—is per annum. As a result, we average the rhetoric score for each group’s documents in a given year. Our explanatory variable of interest, whether or not a group received external financial support, is binary.

We conduct four specifications of a fractional logistic model using **Stata 14.0**. We first estimate a bivariate model of external support’s effect on the rhetoric ratio, then a model that includes the covariates of group casualties and number of fighters. The former is measured in the hundreds of persons, the latter in the thousands of persons. Recall that the STM analysis accounted for the effects of group authorship, year, and language. Yet, as a robustness check, we include a specification using group fixed effects, as well as one including both group and temporal fixed effects (see Appendix F). For the temporal fixed effect specification, we divide the years between 1979 and 2001 into five periods established in the literature on the Afghanistan conflict (Rubin 2002; Christia 2012): the early Jihad, the late Jihad, the intra-*mujahideen* civil war, the early Taliban period, and the late Taliban period. The results are reported as mean marginal effects, derived using the `margins dydx ()` command. These coefficients should be interpreted with no funding (*i.e.*, funding = 0) as the base level. Standard errors are clustered by group and period.

As a further robustness check, we drop one prominent radical group’s observations and re-estimate the fractional logistic models with group and period fixed effects. These results are reported in Appendix F.

Next, we estimate a second set of fractional logistic models using a multiplicative interaction term. Data on the moderator, global oil price, is obtained from the International Monetary Fund’s public database. The IMF reports oil price as a simple average of three price measures—Dated Brent, West Texas Intermediate, and the Dubai Fateh—in US Dollars, indexed to 2005. We interact this moderator with the origin of groups’ external support. External support is labeled as being from one of two places: (1) Saudi Arabia or other predominately Arab states in the Persian Gulf region, oil-rich countries or “petrostates”, and (2) any other country. We identify groups’ patron with data from the UCDP External Support dataset (Högbladh et al. 2011) and by the assumption that groups with large numbers of Arab fighters receive support from Arab Persian Gulf states. This assumption is used to identify four groups as receiving support from the Arab Persian Gulf: the Haqqani Organization, Miramshah College, *Jami’ah al Da’wah ila al-Quir’an wa al-Sunnah*, and *Maktab al-Khidamat*.

Iran, an oil-producing country, also provided external support to some groups in Afghanistan, but we only select groups with ties to the Arab Persian Gulf into the “treated,” or oil-state supported, category because monetary resources in Iran are relatively less sensitive to global oil prices due to sanctions imposed on its petroleum industry and international trade after 1979 (and expanded in 1995). After the imposition of sanctions, Iran’s oil production collapsed from pre-1979 levels, making oil production a smaller part of its economy. In addition, the sanctions meant Iran’s oil production and revenue did not respond to oil price as production and revenue did in the Arab Persian Gulf. For example, during the rise in West Texas Intermediate prices in the early 1980s and the (expected) corresponding increase in production by oil-rich Arab states, Iran’s production decreased.

Because oil price is recorded at monthly intervals, we use month-level observations in the multiplicative interaction portion of the analysis. This means we return to our original rhetoric ratio values—each group

⁴¹<http://ucdp.uu.se>

⁴²<https://icr.ethz.ch/data/epr/acd2epr/>

is assigned a mean rhetoric score at monthly intervals. Doing so gives us more observations and greater statistical power. The disadvantage is that we cannot include the annual measures of group casualties and number of fighters as covariates. However, we use fixed effects, which our initial model results indicate produce coefficients in the same direction and at the same level of significance as the specifications with the covariates (see Table 3 in the main text and Table F6 in Appendix F).

For the multiplicative interaction specifications, we use the `interflex`⁴³ package for R (Hainmueller et al. 2018). The model uses a simple binning estimator. This approach offers several advantages, as reported in section 4.1 of Hainmueller et al. 2018. We highlight two: first, it serves as a formal test of the validity of the linear interaction effect assumption imposed by the standard model (the assumption is validated in our case) and, second, conditional marginal effects are estimated at typical values of the moderator (in our case, the medians of the first, second, and third terciles of the moderator). The dependent variable, the monthly rhetoric ratio, is a continuous value falling between zero and one. The focal independent variable is dichotomous, indicating whether a group received support from an Arab state in the Persian Gulf in a given month. The moderator is the price of oil, measured in U.S. dollars. This last unit of measure is appropriate because marginal effects do not provide a very good approximation of the effect of a one unit increase in the independent variable when the variables unit is large, but they *do* provide accurate approximations when the units are small, such as a dollar.⁴⁴ Standard errors are clustered by group and period. We report the findings as mean marginal effects.

To evaluate the impact of our assumption that groups with many members from the Arab Persian Gulf states receive support from those states, we estimate the fractional logistic multiplicative interaction model when the category of groups supported by the Arab Persian Gulf states only includes, first, groups that explicitly received support from Saudi Arabia (as recorded by UCDP) and, second, only the four groups we assume to have ties to the Arab Persian Gulf based on their high numbers of Arab fighters (*i.e.*, the Haqqani Organization, Miramshah College, *Jami'ah al Da'wah ila al-Quir'an wa al-Sunnah*, and *Maktab al-Khidamat*). This check verifies our identification of support from the Arab Persian Gulf (see Appendix F).

Finally, we check the impact of our assumption that radical backers in Saudi Arabia and the Gulf base their decision to offer support on the wealth they expect to receive from current oil prices. Under this assumption, we do not lag the price of oil. However, because we do not have strong evidence in support of this assumption, we also estimate our model using oil prices lagged by one month and six months. The overall finding remains the same (Appendix F).

⁴³Version 1.0.3; <http://yiqingxu.org/software/interaction/RGuide.html>

⁴⁴<http://www3.nd.edu/~rwilliam/xsoc73994/Margins02.pdf>

E Effect of Support from Saudi Arabia or Arab Gulf States

Table E5 displays the coefficients from the fractional logistic multiplicative interaction models. The first model estimates the effect of support from Saudi Arabia or the Arab Persian Gulf, as well as that of the moderator, global oil price, without the interaction term. We find a significant positive effect for origin of support ($p < 0.05$).

The second model adds the interaction term; its coefficient is positive and significant ($p < 0.01$), as shown in Figure 3 in the main text. The main effect of oil price—or, the effect of oil price on groups that have support from outside SA or the Gulf—lessens the rhetoric of subversion. This might be because of the redistribution of wealth when oil price increases. Namely, residents and governments of non-oil producing states have to spend more money on oil when prices increase, resulting in less support available to radicals. This decreasing availability of (potential) external support produces a corresponding decrease in the rhetoric of subversion, as our argument suggests.

The remaining models serve as robustness checks, as explained in Appendix F; the results are discussed in subsection F.3.

Table E5: Effect of Support from Saudi Arabia or Arab Gulf States on Subversion Rhetoric, Moderated by Oil Price

	(1)	(2)	(3)	(4)	(5)	(6)
Price of Oil	0.000726 (0.00100)	-0.00144** (0.000213)	0.000353 (0.00123)	-0.00157** (0.000245)	-0.000371 (0.000719)	-0.00133** (0.000186)
Support from Saudi or Gulf	0.0809* (0.0391)	-0.0377 (0.0581)				
Oil Price X Support from SA or Gulf		0.00343** (0.00116)				
Support from Saudi Arabia			0.103** (0.0397)	-0.0273 (0.0776)		
Oil Price X Support from SA				0.00367* (0.00172)		
Support from Gulf					0.231** (0.0510)	0.144** (0.0549)
Oil Price X Support from Gulf						0.00253** (0.000891)
Group FEs	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	16,932	16,932	12,754	12,754	7,031	7,031
Pseudo <i>R</i> ²	0.046	0.047	0.022	0.022	0.105	0.105

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Coefficients are mean marginal effects derived from a fractional logit model. Standard errors, in parentheses, are clustered by group and period. Results are based on two-tailed tests.

F Adoption of rhetoric: Robustness checks

F.1 Fractional logistic regression

Our first robustness check entails modeling the bivariate relationship between external support and the rhetoric of subversion using, first, group fixed effects (FE) and, second, group and temporal FE. We estimate this specification as a robustness check because the STM analysis accounts for the effects of group authorship, publication year, and language of the text. Table F6 presents mean marginal effects of the FE specifications, with no support as the base level.

When including group FE, receiving funding is associated with an increase of 0.232 ($p < 0.01$) in our rhetoric ratio. When using both group and temporal FE, the increase is 0.257 ($p < 0.01$). Considering that the mean rhetoric ratio is about 0.65, these increases are substantial. These results mirror the findings presented in the main text.

Table F6: Effect of External Support on Radical Rhetoric of Subversion, with Fixed Effects

	(1) Rhetoric Ratio	(2) Rhetoric Ratio
External support	0.232** (0.0787)	0.257** (0.0904)
Constant	0.545** (0.101)	0.726** (0.0908)
Group FEs	Yes	Yes
Period FEs	No	Yes
Observations	12,802	12,802
Pseudo R^2	0.019	0.021

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Note: Coefficients are mean marginal effects derived from a fractional logit model. Standard errors, in parentheses, are clustered by group and period. Results are based on two-tailed tests.

Our second robustness check entails removing some prominent groups from the dataset. The findings are largely robust to these exclusions (Table F7). Excluding *Jam'iyyat*, *Jabha*, and *Hizb-i Islami* (Gulbuddin faction) does not change the direction and significance of the estimate, and only has a minor impact on the effect size. This check suggests the relationship between external support and rhetoric is not occurring within a specific group. The exception is the Taliban: when Taliban observations are removed, the effect becomes non-significant, although the direction of the effect remains positive. This change might be due to greater variance introduced by removing the relatively large number of observations ($n = 11,181$). However, it is worthwhile to consider that the Taliban might be an exemplary case of our argument—they did not receive as much external support as the *mujahideen* groups did during the *Jihad* (Zaeef 2010).

Third, we seek to gain further confidence in our conclusion that radicals' networks of support affect their use of language by estimating the effect of external support on the prevalence of individual topics. Doing so also sheds light on the extent to which our insights depend on our (theoretically informed) aggregation of topics into rhetorics.

We estimate seemingly unrelated regressions (SUR), or multiple equations related to one another through a correlation in the errors. This way, we capture the fact that resources can have different but linked influence on the prevalence of each topic. In other words, we estimate the effect of resources on the prevalence of a

Table F7: Effect of Support on Rhetoric of Subversion, Excluding Prominent Groups

	(1) No Jam'iyyat	(2) No Jabha	(3) No Hizb-i Islami (G)	(4) No Taliban
External Support	0.204** (0.0658)	0.257** (0.0906)	0.268** (0.0959)	0.245 (0.278)
Constant	0.839** (0.0658)	0.726** (0.0910)	0.715** (0.0963)	0.738** (0.278)
Group FEs	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes
Observations	6,729	12,443	9,877	11,181
Pseudo R^2	0.029	0.020	0.027	0.005

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Each column presents the results of models estimated without observations from one key group. Coefficients are mean marginal effects derived from a fractional logit model. Standard errors, in parentheses, are clustered by group and period. Results are based on two-tailed tests.

given topic, relative to all other topics, while taking into account that the effect may be related to the effect of resources on each of the other topics.

Figure F7 presents the results. We find that receiving external support increases the prevalence of war-like topics. At the same time, groups receiving external support decrease their use of topics forming the rhetoric of reversion. While there is also a decrease in some of the subversion topics, they decrease relatively less in comparison to the reversion topics. Overall, these findings align with the expectations of our argument and main conclusion: receiving external support helps shift radicals' rhetoric towards subversion and away from reversion. However, in light of the evidence of heterogeneous effects across topics, we recommend that future research explores further the relationships between resources and individual topics, as well as between topics and rhetorics.

F.2 Accounting for groups' geographical area of activity

Between 1979 and 2001, most, if not all, of Afghanistans radical groups had representatives of their leadership abroad, in relatively easy to access urban centers, such as the Pakistani cities of Peshawar and Quetta. However, many members of the groups' mid-tier leaders ("commanders") and soldiers lived and operated within Afghanistan, which has large portions of rugged and inhospitable terrain that made travel, transportation, and communication challenging. Consequently, while a group's leadership may have had certain relationships of support, their colleagues in the field could have experienced differing access to resources.

To account for this possibility in our analysis, we re-conducted the moderation analysis with the inclusion of a covariate reflecting groups geographic area of activity on a monthly basis. To create this variable, we drew on data on violent incidents involving groups that were contained in the UCDP dataset⁴⁵ (see Appendix D). If a group was active in an attack located in a region of Afghanistan in a given year (the temporal unit used by UCDP), the group was coded as operating in that region during each month of that year. We apply observations at the annual level to months because insurgent groups are often active in an area before and after a violent incident may be recorded by observers. In addition, fighting in Afghanistan followed (and still largely follows) a seasonal pattern—insurgents established areas of operations during the warmer months and regrouped during winter. We used the commonly known regions of Afghanistan: north, west, central, east, and south. Including the area of operation covariate constrains our analysis to 2,480 monthly observations of groups and their rhetoric. Our main analysis relied on 23,412 monthly observations of groups and their rhetoric.

Our main results are robust to including data on groups' spatiotemporal location. Figure F8 shows that

⁴⁵<http://ucdp.uu.se>

Figure F7: Effect of Support on the Prevalence of Topics

Note: Figure shows the estimated effect of external resources on the prevalence of topics using seemingly unrelated regressions (SUR). The color of the point estimates indicates whether the topic forms part of the subversion or reversion rhetoric. 95 percent confidence intervals are indicated by the black bars.

the effect of support from SA or the Arab Gulf states, moderated by oil price, continues to be positive, and still increases as oil prices rise. The confidence intervals of our estimates are much larger than in our main analysis because of the substantial reduction in observations. Figure F9 shows the results when we only use observations that are missing spatiotemporal data ($n = 17,324$). We obtain findings similar to both the main analysis and the analysis including the spatiotemporal covariate. The two new models presented in this section include group and period fixed effects (see subsection F.1).

F.3 Support from oil-rich Arab states

The third through sixth models presented in Table E5 serve as robustness checks of the main analysis of how support from oil-rich Arab states affects rhetoric, as explained in Appendix D. These models show that interacting origin of support and oil price results in positive significant coefficients, whether considering support only from Saudi Arabia (see also Figure F10) or support only from other Arab Gulf countries. The latter observations are based on the assumption that a high rate of Arab fighters in a group is equivalent to that group receiving support from Saudi Arabia or Arab Gulf states (see subsection 3.3 in the main text). Figure F11 and Figure F12 show that we achieve results similar to that of our main model (Model 2, Table E5) when oil price is lagged by one or six months.

Figure F8: Effect of Gulf Support on Rhetoric, Moderated by Oil Price and Including Group Area of Activity

Note: The dark line and gray 95 percent confidence interval band depicts the conditional marginal effect of support from SA or Arab Gulf states across oil price levels estimated by the standard multiplicative interaction model. The point estimates with 95 percent confidence interval bars represent the conditional marginal effects from a binning estimator at the medians of the three moderator-value bins, the first tercile (L), the second tercile (M), and the third tercile (H). The stacked histogram shows the distribution of the moderator (oil price in U.S. dollars; 2005 = \$100), with the red and white shaded bars indicating the distribution of the moderator in the treatment and control groups, respectively.

The model specifications do not include the covariates of casualties or number of fighters because the rhetoric and oil price observations are monthly, whereas the covariate data are yearly. However, we include fixed effects, which our initial model results show produce coefficients in the same direction and at the same level of significance as the specification with the covariates (see Table 3 in the main text and Table F6 in Appendix F). Finally, note that interpreting the independent effect of support origin—and its coefficients, which are sometimes negative—is substantively meaningless because there are no cases in which oil price equals zero (see Brambor et al. 2006).

Figure F9: Effect of Gulf Support on Rhetoric, Moderated by Oil Price, Using Observations Missing Group Area of Activity

Note: The dark line and gray 95 percent confidence interval band depicts the conditional marginal effect of support from SA or Arab Gulf states across oil price levels estimated by the standard multiplicative interaction model. The point estimates with 95 percent confidence interval bars represent the conditional marginal effects from a binning estimator at the medians of the three moderator-value bins, the first tercile (L), the second tercile (M), and the third tercile (H). The stacked histogram shows the distribution of the moderator (oil price in U.S. dollars; 2005 = \$100), with the red and white shaded bars indicating the distribution of the moderator in the treatment and control groups, respectively.

Figure F10: Effect of Support from Saudi Arabia on Rhetoric of Subversion, Moderated by Oil Price

Note: The dark line and gray 95 percent confidence interval band depicts the conditional marginal effect of support from SA or Arab Gulf states across oil price levels estimated by the standard multiplicative interaction model. The point estimates with 95 percent confidence interval bars represent the conditional marginal effects from a binning estimator at the medians of the three moderator-value bins, the first tercile (L), the second tercile (M), and the third tercile (H). The stacked histogram shows the distribution of the moderator (oil price in U.S. dollars; 2005 = \$100), with the red and white shaded bars indicating the distribution of the moderator in the treatment and control groups, respectively.

Figure F11: Effect of Support on Rhetoric of Subversion, Moderated by Oil Price Lagged One Month

Note: The dark line and gray 95 percent confidence interval band depicts the conditional marginal effect of support from SA or Arab Gulf states across oil price levels estimated by the standard multiplicative interaction model. The point estimates with 95 percent confidence interval bars represent the conditional marginal effects from a binning estimator at the medians of the three moderator-value bins, the first tercile (L), the second tercile (M), and the third tercile (H). The stacked histogram shows the distribution of the moderator (oil price in U.S. dollars; 2005 = \$100), with the red and white shaded bars indicating the distribution of the moderator in the treatment and control groups, respectively.

Figure F12: Effect of Support on Rhetoric of Subversion, Moderated by Oil Price Lagged Six Months

Note: The dark line and gray 95 percent confidence interval band depicts the conditional marginal effect of support from SA or Arab Gulf states across oil price levels estimated by the standard multiplicative interaction model. The point estimates with 95 percent confidence interval bars represent the conditional marginal effects from a binning estimator at the medians of the three moderator-value bins, the first tercile (L), the second tercile (M), and the third tercile (H). The stacked histogram shows the distribution of the moderator (oil price in U.S. dollars; 2005 = \$100), with the red and white shaded bars indicating the distribution of the moderator in the treatment and control groups, respectively.

G Effect of Rhetoric on Future External Support

Does adopting a certain rhetoric affect groups' external support in the future? To provide an initial answer, we first calculated the mean rhetoric for each group by year. Aggregating groups rhetoric into years is necessary because support is observed on a yearly basis. We then use a fractional logistic model, as described in the main text, to analyze, first, the effect of external support on rhetoric and, second, the effect of rhetoric lagged one year on receiving external support. Table G8 presents the results as mean marginal effects.

External support is positively and significantly associated with the rhetoric ratio score. That is, receiving external support increases groups' rhetoric of subversion. This result aligns with the main findings of our study. However, when we lag rhetoric by a year and estimate its effect on external support, we find no significant relationship. This result offers suggestive evidence that as groups increase their rhetoric of subversion, they do not increase their probability of receiving external support in the following year.

Table G8: Effect of Lagged Rhetoric on External Support

	(1) Rhetoric Ratio	(2) External Support
External Support	0.134*** (0.0400)	
Lagged Rhetoric		0.305 (0.576)
Observations	51	41
Pseudo R^2	0.010	0.009

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Note: Coefficients are mean marginal effects derived from a fractional logit model. Standard errors, in parentheses, are clustered by group and period. Results are based on two-tailed tests.