Rethinking Military Strategy:
The Primacy of Exposure in Combat Attrition

by

Dumitru Minzarari

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Political Science)
in The University of Michigan
2019

Doctoral Committee:

Professor James D. Morrow, Chair
Professor Melvyn Levitsky
Assistant Professor Yuri M. Zhukov
Emeritus Professor William Zimmerman

Knowledge, like the sword, is double-edged

Dumitru Minzarari

minzarad@umich.edu

ORCID iD: 0000-0001-7813-1156

© Dumitru Minzarari 2019
All Rights Reserved

Acknowledgements

I owe a tremendous debt of gratitude to my dissertation committee – Jim Morrow, Melvyn Levitsky, Bill Zimmerman and Yuri Zhukov – for their inspiring guidance, thoughtful feedback and generous support. In particular, I greatly value the role of Jim Morrow in shaping my thinking about modern warfare, and in helping me become a more critical thinker.

During my very rewarding experience with the Department of Political Science, I was fortunate to interact with and learn from its exceptional faculty. In addition to Jim Morrow, my thinking on war and conflict, in broader terms, was influenced by the interaction with and work of Scott Page, Skip Lupia, Andrei Markovits, Anna Grzymala-Busse, Ron Inglehart, Allan Stam, and Yuri Zhukov.

I learned a lot, both as a person and as a researcher, from a number of very bright fellow graduate students, including Jean Clipperton, Gary Uzonyi, Vincent Arel-Bundock, Serban Tanase, Paul Poast, Qingjie Zeng, Ali Sanaei, and Andrea Jones-Rooy. They inspired and taught me more than they have realized.

I gratefully acknowledge the generous financial support through a Rackham One-Term Dissertation Fellowship that was instrumental in finishing this project, and the Rackham Graduate Student Research Grant, which allowed me to better understand how policymakers view and think about warfare.

Finally, I am very indebted to my family, for their support. My separate and special appreciation goes to Roos B..

Table of Contents

Acknowledgments ii
List of Figures iv
Abstract v

CHAPTER

I. Introduction and Motivation 1
II. An Examination of the Macrotechnology of Modern Conflict 13
III. Conventional Warfare – the First Day of Somme 85
IV. Insurgency Warfare – Changing the Lenses 149
V. Public Policy Implications 209
VI. Conclusions 228

References 239

List of Figures

Figure 1.1 The costs range driven by time and action 6
Figure 2.1 Firer's lethality and target's vulnerability and exposure 41
Figure 2.2 Conceptual distribution of red and blue forces across battlespaces S_r and S_b 45
Figure 2.3 Probability scheme of belligerent's uncertainty about its opponent's location 47
Figure 2.4 Phase portrait of the system of equations (17): $\dot{x}$=-0.2x-0.4xy; $\dot{y}$=-0.6y-0.4xy 62
Figure 2.5 Phase portrait of the system of equations (17): I quadrant with different ε_i 63
Figure 2.6 Phase portrait of the system of equations (17): extreme eigenvalues and exposure rates 65
Figure 2.7 Phase portrait of the system of equations (17): extreme fire efficiency 66
Figure 2.8 Combat attrition of red and blue forces over time, (27): $\mathcal{E}_b$=.2, $\mathcal{E}r$=.6 68
Figure 2.9 Combat attrition of red and blue forces over time, (27): $\mathcal{E}_b$=.8, $\mathcal{E}r$=.6 71
Figure 2.10 Combat attrition of red and blue forces over time, (27): $\mathcal{E}_b$=.6, $\mathcal{E}r$=.2 72
Figure 2.11 Combat attrition of red and blue forces over time, (27): $\mathcal{E}_b \ll \mathcal{E}_r$ 73
Figure 3.1 Somme River area and the positioning of the British and French forces 99
Figure 3.2 Distribution of British corps at Somme by 1 July 105
Figure 3.3 Topographic map of the combat area at Somme River 107
Figure 3.4 British and German positions at Somme on 1 July 110
Figure 4.1 Two-dimensional counterinsurgency model separating targets and actions 161
Figure 4.2 Conventional war dynamics of positional type (a), maneuver type (b), and insurgent war dynamics (c) 165
Figure 4.3. Schematics of the insurgent's attrition process and variables affecting it 178
Figure 4.4. Monthly deaths by victim, Afghanistan 2004-2009 182

Abstract

How do military capabilities transform into victory on the battlefield? By answering this question, we could get useful insights into, but also be able to inform, the process behind the risky decision to escalate a diplomatic crisis into war. The existing political science literature has viewed the combat conditions that feed this risky decision as static and through the prism of power preponderance. While under certain conditions this assumption is justified, empirical evidence suggests that almost in half of the violent crises it fails to explain conflict outcomes. Moreover, the weakness of power preponderance argument is increasingly accepted by policy experts and practitioners, while trying to replace it through various heuristics.

Besides its policy importance, the question has considerable implications for academic research on war. Most of strategic analysis efforts, employing contest success functions to build strategic interaction models explaining war, have preferred to base them on military capabilities. This is problematic, provided that it does not accurately capture the proper data generating process. A field that extensively studied combat dynamics – the military operations research – has preponderantly used, out of convenience, the military capabilities approach, as well.

I argue that the algorithm guiding how military capabilities transform into victory is determined by the level of combat exposure of troops, and not firepower. I define exposure as being the degree to which a combat force both intersects the set of lethal fire of the opponent and is known by the opponent to be there. By varying the value of exposure, we can directly control the effect of capabilities and their fire accuracy – at low exposure levels these decrease, being able to achieve zero effect. At high exposure levels the impact of capabilities and firepower can manifest their full lethal capacity. I determine these effects by building a combat dynamics model, through a system of nonlinear dynamic equations, which assesses the degree and speed of troops' attrition.

As an important contribution to this model is the theorizing of exposure role in combat, by exploring the intuition of the Shannon information entropy, in the effort to capture the combat uncertainty. I am able to show that exposure, as a manifestation of combat entropy, is indeed the atomic factor, that has primacy in driving combat attrition. Thus, contrary to largely-hold view that power preponderance is essential for determining combat outcomes, I show it is instead completely dependent on exposure. Moreover, I

identify that the number of capabilities only serves as the carrying capacity – the reservoir, from which belligerents draw troops for the fighting engagements.

I illustrate the manifestation of exposure both in conventional and insurgency wars, by examining two main respective cases, while also exploring micro-evidence from a wider range of tertiary cases. I show that exposure effect dominates the effect of capabilities in any type of combat, the difference being in the degree. The logic of the model allows to show that some non-material factors such as morale, do not affect war outcomes the way conventional wisdom poses. By revealing the causal relationship between attrition and exposure and emphasizing the role of the speed of attrition in offsetting the supplies received by insurgent from its supporters, I build a better understanding of both insurgency war dynamics and involved actors' incentives.

CHAPTER I

Introduction and Motivation

Paraphrasing the ancient Chinese strategist Sun Tzu, who famously stated that "all warfare is based on deception" this work obscurely argues that all warfare is based on attrition. More accurately, it specifies that any war outcome depends on the speed of attrition. To understand and reveal these underlying dynamics of war, the analysis identifies, formalizes and shows that it is exposure - a physical property of combat interaction, and not firepower or its accuracy - that guides the speed of attrition specifically and the attrition generally. It finally determines the victory on the battlefield. This is the essential mechanism of war, stripped down to two sentences. Understanding this mechanism was not the main goal – it rather served as a stepping stone to clarify a slightly different question. Nevertheless, this may well be the essential contribution of this work. The next paragraphs present the puzzle that guided this research.

The escalation of a diplomatic crisis into an armed dispute is one of the riskiest decisions that a policymaker can face. Many politicians tend to understand this. Writing about the choice of going to war, Winston Churchill reportedly stated

that after this move the policymaker "is no longer the master of policy but the slave of unforeseeable and uncontrollable events."[1] There are plenty of other examples, revealing how policymakers can either understand the risks of war escalation or fail to fully realize the extent of these risks. For instance, prior to WW I, both Russian Tsar Nicholas II and his war minister did not shy away from accepting that Russia was not prepared for a war. Reacting to French ambassador's remark about Morocco representing a threat to French vital interests, Nicholas reportedly told him to avoid conflicts, as Russia was not prepared for them yet (Wohlforth 1987, 365). In another Russia example, its leaders seemed to have perceived a reduced risk from a potential war with Japan in 1904, which they were willing hence to take, believing the Russian military would win (White 1964, 142). Given this widespread dichotomy, the observed willingness of policymakers to take risks has been perceived as one of the most important predictors of escalation (Morrow, 1987). What does then affect that readiness to embrace the risk of escalating a crisis into war?

It is widely accepted that national leaders' decision to initiate conflicts are influenced by their expectations about both the outcome of war and the ways it can be achieved.[2] This would suggest that they, technically, bet on the probability of

[1] Quoted in Mitzen and Schweller (2011).

[2] Many research traditions tackling this important issue from various angles seem to agree on this, though occasionally invoking differing reasons: conquering and destroying the enemy's armed forces (Clausewitz 1832/1984), due to unacceptable casualties (Singer 1962), costs or anticipated costs (Wright 1942, Richardson 1948, Bueno de Mesquita 1981), decreasing expected gains (Maoz, 1989), achieving initial war objective (Blainey 1988), or among other things, due to domestic politics constraints (Bueno de Mesquita and Siverson 1995, Arena 2008). Factors such as imbalance of resolve (Mack 1975, Rosen 1972, Maoz 1983), risk propensity (McDermott 1998, Haas 2001, Levy 2003), better military strategy (Arreguin-Toft 2001, Biddle 2004, Stam

victory or defeat. In fact, Wagner (1984, 604) emphasizes that, when he insisted that "the way in which decision makers translate the distribution of such commonly known assets of states as population and wealth into probability estimates of the outcomes of various possible wars may be difficult to discern and quite variable".

It is no surprise then that Schelling (1966, 93) indicated that any violent conflict contains so much uncertainty that its outcome is difficult to predict. Uncertainty about wars' outcomes is further greatly aggravated by the fact that these are inefficient *ex post* (Fearon 1995, 383), or in other words – they are costly endeavors. Among other things, this would indicate that in addition to the risk of losing the war, even a victory may feel like a defeat, if the winner enquired costs exceeding her valuation of the good contested through war.[3]

It ensures then, that by improving our ability to understand what affects war outcomes we will increase our effectiveness in contributing to the quality of decision-making in crises and reducing unjustified risk-taking. It is that war-related uncertainty that generates misperception and miscalculation, leading to rational misinterpretation of opponent's signals and responses that are counterproductive (Morrow, 1989). In attempting to address these challenges, the scholarly literature has focused predominantly on strategic interactions, based on behavioral traits and

1996), the value of stakes (Sullivan 2007), regime type (Lake 1992, Reiter and Stam 1998) or stronger capabilities (Organski and Kugler, 1980) have been considered to explain ways to achieve victory or defeat.

[3] Schelling (1966) talked about the power to hurt being bargaining power, and the power to bear costs in return generating impact on war outcome. While belligerents acquire information about battlefield-related costs, they may not be able to stop fighting when the war costs equal their valuation of the contested good, due to this information-collection time lag. An example is the first day of the Somme Offensive, when it took a few days for the British to understand the magnitude of their losses. History knows many cases when victory was viewed as equivalent to defeat due to its high costs, even assigning a label to it – the Pyrrhic victory.

factors incentivizing this behavior. Slantchev (2003, 127) discussed the role of belligerents' expectations about their ability to inflict maximum costs on the opponent while limiting the opponent's ability to inflict costs in return. Advancing a seemingly comparable argument, Sullivan (2007, 520) posed that actors select themselves into wars, when they believe they have sufficient tolerance for costs to attain their war objectives.

The prevalent research approaches included examining the effects of mobilizing war-related resources and the ways they could transfer into victory, favoring the preponderance of force argument. Or, as Jia and Skaperdas (2012) aptly put it, "performance in warfare – in the sense of probabilities of winning and losing – might be a noisy function of military capacities". The most frequently used such "noisy function" is the so-called lottery contest success function. It takes as inputs the fighting efforts of the two sides, usually military capabilities, and generates as outputs the degrees of success achieved, proportional to their force levels inputs (Slantchev 2011).

While these approaches tremendously advanced our thinking about strategic interactions related to war, one of their frequent criticisms points out that there is a significant subset of wars, in which preponderance of force failed to accurately determine the outcome. As Bas and Shub (2016) observed, indicators such as balance of power or dyadic relative capabilities are imprecise and theoretically inappropriate; they therefore posed that there are many inconsistencies in the theoretical foundations of outcome uncertainty. In other words, we are still not fully

understanding the dynamic mechanism through which capabilities transform into war outcomes. Similarly, in the Slantchev and Sullivan examples above, we understand that a race in mutual cost generation would be a strong strategic incentive, though we are unclear how exactly these costs can be strategically[4] generated. A possible operationalization of this insight is to consider that the process of warfighting itself generates additional rational incentives to adjust the belligerents' preferences for the outcomes (Morrow 1985).

Policymakers cannot be aware in advance neither of the level of these future costs, generated by war, nor of these costs' nature.[5] To paraphrase Gartzke (1999) "war is in the error term" not only in the case of war onset, but also in the case of war process, and consequently – the outcome. Due to this uncertainty, leaders may find themselves in a state of the world where they cannot decide whether escalation into war is more or less attractive than the status quo. A conceptual illustration of this intuition is provided in figure 1.1. Ideally, states would like to achieve their war-contested outcome *X*, at time t_0, paying no cost. However, as fighting develops, and the direct costs of war are mounting (personnel, hardware, and supplies), the war will end, in terms of costs, somewhere in the limits of the red dotted box, where

[4] Creating asymmetrical advantage.

[5] For instance, while debating the costs of the Russian military campaign in Syria, the Western experts frequently provided quite elevated costs. In contrasts, the Russian experts pointed out that the sorties of the Russian air force are conducted at the expense of Russian regular military exercises for the year; and most of the ammunition used by the Russian air force was of old, Soviet type and if not used in Syria would have had to be disposed of through costly procedures. See BBC Russian Service, "Skol'ko Stoit Rossijskaja Voennaja Operatzia v Sirii?" [What's the Cost of Russian Military Operation in Syria?] October 6, 2015, accessed at https://www.bbc.com/russian/business/2015/10/151006_russia_syria_military_costs. Also, see Vladymir Voronov. "Voyna na Prodazhu." [War for sale] Radio Free Europe/Radio Liberty (Russian site), March 1, 2017, accessed at https://www.svoboda.org/a/28341611.html.

costs can range from X-C_{i1} to X+C_{i2}. If the war is cost-intensive, like WWI for instance, the t_{n-k} point on the time line would be reached quicker than if the intensity of war was lower, like in some civil wars. This realization would guide the incentives of actors to push the level of costs of the opponent in the range (X, X+C_{i2}), while make efforts to stay themselves in the range (X-C_{i1}, X) when at t_{n-k}. It is a logical manifestation of the insightful realization by Epstein (1985) that belligerents often try to trade space for time, meaning they would retreat from combat after some threshold of losses is reached, or maneuver to diminish their combat losses. From a rationalist perspective, if a belligerent is pushed beyond X and realizes it, she should seek peace or withdraw from war.

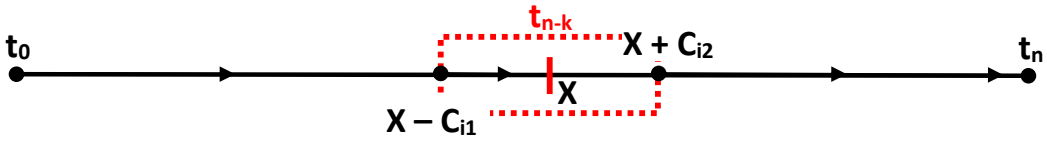


Fig.1.1: The costs range driven by time and action

It becomes increasingly interesting and important then to understand what drives the speed of these costs' accumulation and how one belligerent can achieve advantage over the other in that costs' race. Bringing clarity to this would allow us to better understand the dynamics guiding war outcomes. It is this particular and, arguably, less exploited by political science research niche, of how war efforts

(military capabilities) transform into both victory and costs on the battlefield, that this thesis has the ambition to explore. By getting a better grasp of the related combat ecology and these social mechanisms related to war, the hope is they could more effectively inform our understanding of crisis management decisions, including the risky choice of escalating disputes into war.

We understand quite well the strategic interaction around war initiation and termination in terms of behavioral and motivational factors.[6] This research is contributing to the conflict escalation framework by focusing on the spatial constraints of combat ecology that shape the policymaker's behavioral and motivational factors, following their exposure to that ecology. This approach connects the political science research on crisis escalation, preponderantly focused on behavioral and motivational factors, with the military operations research, focused mostly on the dynamics of combat environment.

It is a natural step, considering the existing research on political preference formation. This thread of literature argues that in order to understand how preferences (including for war escalation) are formed, one needs to explain how beliefs and attitudes are created by the interaction of human observation and feelings (Druckman and Lupia, 2000). According to this research, beliefs emerge as a result of the interaction between the brain, the body and the world (environment). While the existing rational choice literature on war considers directly the brain and the body, through employing the utility maximization conditions, it has paid less

[6] These are qualities like uncertainty, risk-taking propensity, various incentives.

attention to explicitly consider the effects of the world - the physical environment - on decision-making.

The examination of the war dynamics and its effects upon decision-makers' risk-taking propensities that lead to war escalation is nothing else than exploring the "external elements of reason" operating as "incentive-altering forces outside the body" (Lupia et al. 2000, 12). Without considering the relevant effects of the environment – of combat interaction – on policymakers, the exploration of the effects of their "brain and body" on conflict escalation mechanism will generate incomplete insights, at least.

In this research I develop a novel theory[7] about the dynamics of violent conflict, and how military capabilities map onto victory. I argue that the key factor driving the probability of victory is not the capabilities-firepower factor or its accuracy, but the relative degree of lethal exposure of one's forces, which essentially affects these forces' attrition speed. Based on my model of dynamic combat attrition, I derive the intuition that capabilities play the role of nothing more than a reservoir, from which a side can supply the combat efforts. A side with higher number of forces will just have a larger reservoir. Manipulating the degree of exposure requires superior skills in tactics and strategy, or superiority in technology.

[7] I was challenged by a non-academic reviewer as to how novel my contribution is, forcing me to rethink and review this issue. The way I perceive it, is that in the social domain, people have intuitive (more or less accurate) understanding about many things and they tend to use heuristics in that regard that are not that infrequently of spurious nature. They observe developments and try to discern patterns. It is very seldom though when people are able to discern the accurate underlying dynamics of social processes – this is where science comes to rescue. I perceive the novelty of my theory reflected in the fact that it clearly reveals certain traits of an important social dynamics that were not clarified earlier to that extent. And, if accepted, these traits would lead to public policy changes.

The effects of exposure, therefore, are less visible through the human perception. When they are emerging, there is the tendency to attribute them to the respective context. This is why capabilities, strategy, or technology, as more visible to the human eye, have been long selected as some of the dominant explanations for war outcome, while in fact they are only proxies.

This is due to the fact that the degree of exposure of forces can completely block the impact of capabilities and fire accuracy on troops attrition speed, if we decrease exposure through manipulation. It is only when the value of exposure is very high, like in hand-to-hand combat, that the role of capabilities and fire accuracy is maximized. Therefore, it is the smaller exposure of one's military forces that has the greatest impact on war outcome. This way, even a belligerent with considerably lower capabilities can win the war, if she manages to get an advantage in diminishing her troops lethal exposure, comparing to the opponent.

This theory is also able to explain the continuous and non-monotonic dynamics of costs-of-war generation. The variation in the combat exposure of troops can affect the speed of attrition, which first affects the duration of war, and then affects the costs of war. For instance, high exposure of one belligerent allows the other to succeed by spending less munitions, time, and less of her own troops in winning the combat interaction. The significance of my research is that it offers a strong motivation for questioning the dominant, both in the academia and in the policy world, perception of the role of capabilities on victory. The exposure concept that is derived can also explain the effects on combat outcome of non-material

factors such as morale or resolve. Consequently, similarly to the case of capabilities, it can determine the scope conditions of these effects.

In the second and third chapters I apply the logic of combat exposure to both conventional and insurgency types of warfare. I illustrate the role of exposure and its derivatives such as speed of attrition and costs of war, by exploring the case of the first day of the Somme Offensive in WWI. Evaluating the dynamics of the trench warfare and examining the variation in tactics across different units on the battlefield, I am able to gather evidence supporting the essential role of the exposure variable and its primacy, comparing to the preponderance of force factor. The effects of superior firepower were indeed mitigated by the ability of some British and French units to diminish their degree of exposure during attack. In the example of the majority of British divisions, even though they had superior capabilities and firepower, they suffered a disproportionally higher casualty rate due to extremely large exposure. The British divisions to the north and center of their attacking front were technically defeated by the German forces, which, though inferior in numbers, had managed to maintain a significantly lower degree of exposure for their troops.

By exploring the effect of exposure on the degree and speed of troops attrition, I am able to show that the role of morale in conventional combat is unidimensional – higher morale does not necessarily have an effect on probability of victory, independent of the level of exposure. If anything, it could actually lead to higher casualties by affecting the degree of exposure of troops to lethal fire of the opponent. Lower morale, though, could create conditions for increasing the

probability of defeat, as it may decrease the capabilities of a belligerent, while keeping constant the level of exposure. Given that the effect of capabilities on combat outcome is manifested through the variation in exposure, the low morale effect may be considerable. At the same time, low morale can, in certain conditions, decrease the level of troops exposure, allowing to diminish their level of attrition.

While examining the exposure variable in the context of insurgency warfare, I am able to theorize and generate more support for the enemy-centric strategy, in contrast to the population-centric strategy of counterinsurgency. In fact, the effect of exposure is even more pronounced in the insurgency context, provided the asymmetry in its endowment for the two belligerents. By exploring the differences in the Russia's two Chechen campaigns and thus varying the degree of exposure for both Russian forces and the Chechen rebels, it is possible to illustrate the critical and primary role that exposure plays in contrast to capabilities and fire accuracy.

Based on the developed model and the effects of exposure, which operates as an atomic variable, it is possible to argue that coercive strategies can be very effective at the tactical and operational level against an insurgent opponent. These effects are independent of political costs, meaning if the latter are accepted, the coercive strategies will continue to be effective against the insurgent. When the political costs at the strategic level are low or mitigated, the counterinsurgent will be able to maintain a high enough speed of attrition to offset the support that the insurgent is receiving from its local and foreign supporters. The Chechen case suggested that, in addition to the dominant in the counterinsurgency literature idea

that indiscriminate violence is a substitute for intelligence, it can also diminish the exposure of the counterinsurgent forces, balancing the odds in its competition against the insurgent.

In the next chapter I develop and present the theoretical foundations for my argument, applying mathematical modeling to improve the intuition about combat dynamics and propose a causal link between exposure and attrition. Then, in the following chapters, I sequentially explore case studies and micro-comparative evidence to illustrate the manifestation of my theoretical propositions, first in a conventional war and then in the context of insurgency war. I am also exploring the public policy implications of my theory, before concluding.

CHAPTER II

An Examination of the Macrotechnology of Modern Conflict

Introducing the puzzle

On November 19, 1990, the NATO and Warsaw Pact member countries signed one of the most comprehensive conventional arms control treaties in history – the Treaty on Conventional Armed Forces in Europe – also labeled the "cornerstone of European security"[8]. The treaty limited the deployment in concrete geographic areas of specific equipment for both parties, believed to be of "assault" nature[9].

Before World War II the French military invested significant resources to build at its eastern borders a series of fortifications that were labeled the Maginot Line (Kraehe 1944). They were meant to deter a surprise attack by Germany, or force the German troops move around the fortifications to an area where they were expected. Similarly, the Finnish Government built during 1920-1930s the

[8] See the interview with Ambassador Thomas Graham, President Clinton's Special Representative for Arms Control, Non-Proliferation, and Disarmament (Mendelsohn and Walkling 1996). For a detailed examination of the conditions that led to the emergence of this treaty, see Falkenrath (1995).
[9] This included battle tanks, armored combat vehicles, artillery, combat aircraft and attack helicopters.

Mannerheim Line – a range of fortifications on the Karelian Isthmus intended to help Finland protect against a possible Soviet invasion (Irincheev 2009).

In a more recent historical development, the Taliban envoy to Pakistan Abdul Salam Zaeef refused, on behalf of the Afghan Taliban Government, to hand Osama bin Laden over to the United States, which was requested through an ultimatum by U.S. President George W. Bush. Zaeef responded that "his government was ready, if necessary, to defend the country against American attack"[10].

One main thing that all previous examples have in common, I argue, is that they represent instances of strategic calculations by political actors, which are based on certain assumptions about how military capabilities could transfer into military victory. The most likely way we could *rationally* explain the actors' readiness to escalate a diplomatic dispute into violent conflict is if the decision-makers had specific believes about how the military tool could bring victory on the battlefield (Fearon 1995, 386).[11] Therefore, these assumptions are essential to the decision of policymakers of either escalating or not a crisis from diplomatic bargaining into armed conflict.

It is this important question that the present analysis is going to explore - how does military power transform into victory on the battlefield[12] and what affects the

[10] WIRED. *Taliban Rejects the U.S. Ultimatum*. September 21, 2001. https://www.wired.com/2001/09/taliban-rejects-u-s-ultimatum

[11] Even if these decision-makers perceived initially their opponents were not resolute enough to attack, then following the attack there must be plenty of opportunities to engage the attacker in negotiations that would offer a better bargain than the one expected to emerge after war. Some literature on war bargaining allows for negotiations to continue during the fighting (i.e. Filson and Werner 2002).

[12] The question could be paraphrased to ask: "why does power preponderance sometimes translate into military victory and sometimes not?" However, the focus of the question is specifically on the dynamics and mechanisms

variation in the observed outcomes? Besides its obvious policy relevance, bringing some clarity to this puzzle has considerable benefits for political science research. I argue that this analysis makes a sensible contribution to political science literature that explores crisis escalation.

One main contribution is that it bridges the two bodies of literature on conflict that speak past each other: the bargaining diplomacy scholarship in political science and the military effectiveness literature of the military operations research. Another contribution of this analysis is the introduction of a different optics on military effectiveness. It suggests a dramatic shift from preponderance of force and firing accuracy thinking towards the ability to avoid targeting through the variation of troops exposure to lethal fire. This advances a rather different perspective because the leading relevant political science (and military operations) research has been built on the power preponderance assumption. Among other things this analysis has the potential to clarify the conditions mitigating the disagreements between the capability model and the resolve model in explaining disputes' outcomes (Maoz 1983) and thus reconcile them.

The concept of combat exposure will be introduced analytically through the assessment of combat microdynamics. I do this by departing from the analytic convenience of distinguishing between interstate and intrastate violent conflicts, which allows me to conduct a richer exploration of combat dynamics. It changes the

explaining how this happens. I argue that only by first understanding the underlying dynamic of a social process (the "how?") can we get a best shot at understanding why that process occurs.

combat outcomes distribution by adding cases to the subset of conflicts in which weaker sides were able to successfully confront a more powerful opponent. This is also a novel endeavor in political science research on conflict, which treated combat microdynamics like a black box by making largely heuristic assumptions about it. Consequently, exploring the ideas of this analysis will bring more clarity about the military effectiveness assumptions that are foundational for various models of crisis escalation. It will suggest new relevant variables to be considered in the empirical analysis of conflict. For instance, the presented dynamic model of conflict suggests that troops' exposure to lethal fire and the speed of attrition of fighting forces have a higher impact on combat outcome than the sheer number of capabilities or their firing effectiveness. The analysis also generates insights into how exposure factor affects the costs of war, both directly and indirectly. Finally, I argue that by offering new insights into the war dynamics and what impacts probability of victory in war, as well as the costs of war, we could further improve our overall understanding of crisis escalation dynamics.

In contrast, the dominant approach in the literature has been to generally assume that the chances of victory in a violent conflict are proportional to the quantity of military capabilities of the competitors. This view has transcended across various paradigms and intellectual traditions of the discipline. The respective literature is voluminous, but a few examples of influential work would help make the point.

The realist tradition of international relations has famously explored the concept of national power in terms of aggregating military capabilities and strategic resources that support them, including population and natural resources (Morgenthau 1948, Singer et al. 1972, Organski and Kugler 1980). Some of the most used war-research data sets, which guided intellectual inquiries of war for decades, took a similar approach (Singer 1988). One of the most influential traditions of crisis escalation research employing the bargaining model of war has regularly been assuming that preponderance of military force is the best indicator in determining chances of victory in war (Reiter 2003).

All these approaches have brought valuable insights in advancing the understanding of crisis escalation. They have both historical and public policy rationale, which justifies their specific choices. For instance, a persistent debate in the US military has been whether quality or quantity of military force is most important for achieving victory in war (Lepingwell 1987, 89). Historical studies, which considerably influenced political science research, have also emphasized the importance of perceived relative strength of belligerents in war (Blainey 1973). A related and relevant thread of debate includes the role of technology, as a multiplier of military capabilities, for ensuring war victory. The political science research examining the balance between offense and defense and its effect on crisis escalation is a good illustration in this regard (Van Evera 1998).

It was not until more recently, in historical terms, that both researchers and practitioners started to observe some empirical contradictions to that long-held view.

A number of failures by militarily-dominant forces to ensure victory against weaker opponents, either state or non-state, had casted doubts on the essential role of force preponderance. Failures of France in Indochina, of USSR in Afghanistan, and United States in Iraq and Afghanistan more recently are just some of the usually-quoted examples.

Military professionals, policymakers, and academic researchers started to increasingly accept the idea that in modern times superior capabilities do not necessarily predict victory[13]. For instance, Helmbold (1964, 781), after assessing Lanchester equations through a statistical examination of 92 historical battles contended his analysis had revealed that "victory in battle is primarily determined by factors other than numerical superiority, and challenges the ability of any model of combat which concentrates almost exclusively on numerical force size to yield a practically useful predictor of victory in battle". Similarly, Stam (1996, 73) suggested that "forecasts based on simple measures of capabilities, are crude, at best".

Attempting to explain the failure of power preponderance to assure victory in war analysts explored the impact of strategy and force employment on military victory (Stam 1996, Arreguin-Toft 2001, Biddle 2004), of resolve (Rosen 1972, Mack 1975, Mueller 1980, Maoz 1983), geography (Stam 1996, Fearon and Laitin 2003) or cost tolerance (Sullivan 2007, Slantchev 2003). It is important to notice that some of the relevant influential research confounded the effects of capability and resolve factors (Morrow 1985, Powell 1987, Fearon 1995). Some of more resent

[13] Interviews conducted in Brussels in the summer of 2014, with European military and civilian officials.

investigations claimed resolve was independent of military capabilities and instead was purely psychological (Kertzer, 2016). This suggests that understanding how capabilities transfer into victory may be useful in shedding light on other, logically neighboring concepts and debates, including resolve.

The listed policy and academic debates would indicate that there is a knowledge gap in existing research, noticeable in particular because this fact began to influence the thinking of policy practitioners. Given that the power preponderance model is not able to offer satisfactory explanations to the increasing number of empirical contradictions listed above, there clearly must be some other factors at play. The current analysis is focusing on identifying and shedding light on these additional factors.

A starting step in identifying factors other than power preponderance that are affecting combat outcome is to understand the causal mechanism at play. The variable of war outcome has been the subject to considerable statistical analysis, exploring correlation links with various factors, including capabilities. Nevertheless, there was hardly any effective attempt to explain the specific causal mechanism of how war outcome is ensured. Although military operations research did examine various factors affecting war outcomes, which have been used by modern Western armies to improve their war-fighting abilities, political science research avoided delving thoroughly into the war macro and microdynamics question. It is despite the fact that this subject continues to remain a very important one for military and

strategic policy planners, powerfully affecting the policymaking on crisis escalation. Therefore, it presents a significant interest for political science too.

If we want political science research to be viewed by policy-makers as relevant, useful, and be accepted, we should strive to incorporate and reflect in our research the practical preoccupations and needs of the policy practitioners (Desch 2019).[14] This also responds to one of the key expected values of scientific endeavor – the provision of public good – by facilitating the transfer of academic knowledge into effective public policy. It is an important objective of this analysis to address this concern, in addition to the goal of contributing to the crisis escalation academic literature.

In the next section the analysis will explore the current political science research efforts on war dynamics. Following that it will examine existing approaches to formal modeling of war dynamics in the technical literature, including military operations research and economics. Based on the identified gaps, the analysis will then conceptualize and introduce a new global primitive significantly affecting the macro-dynamics of war – the troops exposure to lethal fire. This variable will be explained through the formalization of combatants' uncertainty about the opponent, using the information entropy. A more detailed explanation of the microdynamics of war will be offered by presenting the clarifying context necessary for the conceptualization of exposure. Advancing these considerations, the analysis will

[14] Desch suggested that, according to polls, senor national security decision-makers, overall, believe that academic social science is not offering them useful insights to work with.

propose an adapted model of war, exploring the non-linear dynamics of combat interaction trough ordinary differential equations. Building on this war model, the analysis will present its second contribution by emphasizing the importance of the speed of attrition for both military and political components of war victory. Finally, the conclusion will analyze the model results and discuss their significance.

War dynamics in conflict research

Before presenting my own analysis, I would like to examine the respective state of affairs in political science. This should allow clarifying how the current research would differ from, but also would further advance the existing research endeavors by contributing to existing academic debates and efforts. It would also allow revealing how, generally, could political science further contribute towards improving our understanding of crisis escalation mechanisms and processes in ways useful to public policy.

A large part of the research on crisis escalation is focusing on how the decision to either escalate a dispute into war or yield to pressure is shaped by the competing sides' perceptions of their ability to inflict mutual harm upon each other. And this ability is considered dependent on military power. One of the most quoted analyses explaining war initiation, exploring the bargaining model of war, suggests that uncertainty about the capabilities or resolve of the opponent may lead to war (Fearon, 1995). This conclusion though, as well as many similar others is dependent

on the model specifications selected by the researcher. That model explicitly uses a capability proxy variable to represent the probability of victory. It poses that the side with stronger capabilities is more likely to win the war. This approach is rather the rule than the exception in modeling war escalation in political science (Bueno de Mesquita 1981, Powell 2006, Slantchev 2011).

This modeling protocol choice has brought valuable insights into crisis escalation research. It is a technically simple and a logically defendable way to represent how a specific resource input generates a specific political output, similarly to the production function in economics. In an analogy to production function, in conceptualizing violent escalation dynamics, political scientists referred to that input-output process as the technology of coercion (Powell 1999, 14) or technology of conflict (Slantchev 2011, 81). As Powell explicitly puts it, "the greater an actor's coercive capabilities, the more powerful it is". Slantchev suggests that "the probability of victory depends on one's share of the total military forces". More recently, the technology of coercion has been referred to as contest success function (CSF) - a term that this analysis is going to use as well. It reflects more thoroughly and formally the Powell's explanation, suggesting that it receives the war-fighting efforts of the contenders as inputs and produces as output a probability of victory (Hirshleifer 2000, 774).

There is nothing wrong about the construction of CSF based on power preponderance as such is the nature of the modeling enterprise – to make simplifying assumptions that help acquire useful analytical insights. What is

important to notice though, is the need for a more rigorous understanding of combat dynamics. This would allow us to generate more accurate underlying assumptions for the CSFs, leading to a higher fidelity in reflecting how military power transforms into probability of victory in war. Or, without this the conclusion will be restricted to only reflect the chosen structure of the CSF – if it is capability-based, then given its probabilistic nature the conclusion will invoke some form of uncertainty about capabilities. If we replace the capability and resolve variables with something else, the CSF will respectively generate an output about the new variable. Therefore, we shall build the respective CSFs on a more accurate understanding of war dynamics. Any attempt to explain crisis escalation should apply a theoretically-sound and empirically-grounded formulation of the CSFs that determines the crisis escalation model, either the bargaining one or of any other type.

As already observed, there have been a number of important recent empirical cases challenging the capability-based choice of the conflict success function, at least partially. These imply that under certain conditions, greater coercive capabilities do not necessarily transform into greater power and consequently into military victory. Arreguin-Toft (2001, 97) suggested that since 1800 till 1998, the share of weaker side's victory in the total number of violent conflicts has increased constantly from 11.8% to 55%. Following the period analyzed by him the trend has persisted. A similar empirical regularity has been captured in the work of other researchers, suggesting that in over 40% of wars during the last 200 years the stronger states were not able to prevail (Bueno de Mesquita, 2000). Provided

influential research on war escalation suggesting war is very risky and costly even for powerful contenders (Schelling 1966, 93), it is puzzling when a weaker actor resists or challenges more powerful contenders and consciously risks war. If capabilities were the only, or even the dominant factor, it can be argued that we should see more victories by stronger belligerents and less instances of weaker sides challenging the more powerful.

Of course, the results above depend to a large extent on the conflict data-coding choices[15] and whether one separates interstate conflicts from conflicts involving state and non-state actors. If we add to these data the armed conflicts at intra-state level, the failure rate of the stronger contender to win dramatically increases. There are strong reasons to do this. The war dynamics, as it will be explained in detail further in the text, is identical in both types of conflicts. A separation of conflicts into interstate and those between the state and non-state actors is more a subjective analytical choice and a matter of convenience. It is not necessary if our main focus is on the dynamics of warfare. In fact, military professionals and analysts view the underlying dynamic of any armed combat as being one of attrition[16] in which war-related resources such as troops, equipment and logistics are being degraded during combat interactions (Ancker 1995, Perry and Moffat 1997). And the attrition is characteristic to both conventional and

[15] See Reiter et al. (2016) for a criticism of the COW data set and its coding choices.

[16] There is a frequent practice in the political science literature to view attrition (warfare) as one of the "operational strategies" (Malkasian 2002, 4) along maneuver strategy and punishment strategy (Stam 1996, 52-53). For the purpose of my research I do not see analytical benefit in this classification. In particular, since military practitioners themselves do not explore it, which means it has less effect on public policy. Besides, as I explain in the text attrition is a dynamic process that describes the degradation of war-related resources in any combat interaction – a concept largely explored by military operations research.

unconventional (insurgency-based) military conflicts. As Hirshleifer (2000) put it in his illuminating analysis on conflict, “at the point of contact, all clashes come down to attrition”. The difference between various types of warfare, as this research will show, is in the speed of attrition. These are some additional reasons why I suggest avoiding a conceptual difference between interstate wars and wars between state and non-state actors. Besides, logic suggests that these extreme cases in which the weaker challenges the stronger may better reveal what are the driving forces behind war escalation.

Given data on conflict outcomes, then, statistically speaking, by persisting to use the capability-based approach to assess probability of victory we are throwing away a very large part of our data by excluding some violent conflicts. As a result of this, we do not capture accurately the relevant data generating process. And, this cannot go without impact on the accuracy of the produced conclusions, generated by various conflict escalation analyses, including those implied by the bargaining model.

To avoid these negative effects, I would like to argue and suggest that it is of critical importance to more thoroughly identify the assumptions that feed into the conflict success functions of the conflict escalation models. Some attempts in that direction have already been made. Besides the conflict success functions, reflecting the simple distribution of power of the conflicting sides (Slantchev, 2011), researchers produced more recently CSFs that attempted to capture the amount of resources and efforts invested by the belligerents into the combat activity. It can be

argued though that to a large extent this also ends up reflecting the logic of distribution of power approach. According to these other CSFs, the side producing the highest fighting intensity would be more likely to win. Or, given a similar fighting intensity, the side with higher initial amount of war-related resources would win (Hirshleifer, 781). This type of CSFs is seemingly accommodating the arguments about resolve and capabilities through firepower as well as fire accuracy.

Regardless of their shortcomings, the use of these more elaborate CSFs in the analysis of conflict escalation, including the bargaining model, would be a step forward for the discipline. However, provided that military professionals have been using some of them already for decades, they will bring little to the general understanding of war dynamics itself. They could, nevertheless, enrich our understanding of the political repercussions triggered by a more sophisticated view on war dynamics. It has to be stressed, though, that the exposure concept which I propose to use for adjusting CSFs may also be useful to the military professionals. It offers a different perspective on war attrition given the military focus is preponderantly on capabilities and firepower effectiveness.

What is the policy relevance of accepting my proposed adjustment to existing research approaches on crisis escalation? To clarify this better, let us explore two related policy scenarios that illustrate the possibility of improving our understanding of crisis escalation mechanisms and processes. In a first story, the policymaker, who decides whether to escalate or not a diplomatic confrontation into violent conflict, is following during crisis escalation a strategic logic that is somewhat

different from the one suggested by current political science research.[17] Regardless whether this is right or wrong, it should be noted that different thought processes in policymaking could generate different actions (Barabas 2004) and thus outcomes. One way to help improve public policies on crisis escalation would be to make the existing influential research more actionable and applicable by policymakers. For instance, we could ask ourselves whether policymakers, during their crisis escalation involvement, think in terms of asymmetric information and readiness to make deals if the missing information would become common knowledge.

According to a second story, let us suppose that the policymaker acquired an erroneous picture compared to reality on how exactly military capabilities transform into victory. Because of this, she is likely to make wrong cost-benefit calculations about escalating a crisis into war. To generate added value in terms of public policy, political science would have to identify the error and suggest policy-actionable solutions. In practical terms, we know that a policy decision or conclusion can be the artefact of a particular political process or of some specific political interests (Drezner 2000, Heisenberg 2005, Nowlin 2011). Therefore, we could instead directly examine the war dynamics, and enquire whether a (currently dominant) capability-based view on war dynamics is accurate and helpful.

Another question to ask then, is if by replacing a capability-based model with a troops' exposure model, we could trigger different conclusions through the use of

[17] Interviews with security and military practitioners at NATO and EU suggested this is the reality. To be more accurate, nobody from those interviewed was aware of the existing rationalist explanations of war examining asymmetries of information and commitment problems. When introduced to the concepts they expressed uncertainty about how these could be instrumentalized in policymaking.

the bargaining model of war. For instance, would a change in our understanding of combat dynamics make a difference in the ways policymakers assess risk? If the answer is positive, then changing the way we view war dynamics is able to bring added value to policymaking. This chapter is focusing on that latter scenario.

So far, a concrete suggestion is crystalizing. To be able to make better-informed and even more accurate cost-benefit calculations involving the escalation of a crisis into war, a policymaker would need to have a decently accurate understanding of war dynamics. Or, more technically, she would have to understand how her country's military capabilities and other related traits could transfer into victory on the battlefield. If policymakers do not know this, and we have significant reasons to believe that they do not (O'Hanlon, 2003), then the decision to escalate into war may be the result of optimistic expectations about war and its related costs (Slantchev and Tarar, 2011). Or, viewed from a parallel perspective, as thoughtfully argued by Garztke (1999), policymakers' perception and decisions on crisis escalation into war would be affected by the error term, triggered by high uncertainty about war dynamics. However, even in this case, knowing what affects military victory is extremely useful to policymaking and risk assessment.

To aggregate this logic, if competitors do not have effective knowledge about how the interaction of their capabilities or of their resolves can produce victory, then they *a priori* cannot make accurate policy-relevant inferences about the effect of capabilities or resolve on victory. As Fearon (1995, 386) has aptly put it, states must be able to evaluate military options in comparison to potential outcomes in order to

be able to know if there are negotiable settlements available. This echoes well with Biddle (2004, 27) argument that in order to understand the effects of preponderance of power one must look at force employment at both operational and tactical levels. He then suggests that the military-strategic level may not be able to clarify the effects of capabilities in combat to the same extent as the operational and tactical levels (Biddle 2004, 253).

The problem is that neither practitioners nor war theorists seem to have a solid grasp on why exactly sometimes superior capabilities are a predictor of victory, while in many other instances they are not. The existing perceptions are based on observations and interpretations that lack a strong conceptual ground. Theoretically, we do not have a clear-cut idea of a CSF that would describe how the war effort maps onto victory, with a decent degree of fidelity.[18] To use a more game-theoretic terminology, it is likely that the belligerents do not have a common conjecture about the reasons why power preponderance does not always transform into victory. And if parties to the conflict have different views on how capabilities may affect war outcomes, this could create grounds for escalating a conflict into war.

In the next section, I will discuss the existing attempts to model conflict dynamics and their shortcomings. This shall prepare the ground for the discussion of my own approach of assessing what affects the probability of victory and for the proposal of an alternative model of war dynamics.

[18] This is not a critique of the probability-based assessment of war outcome. Instead, it is a suggestion that by improving the conceptual foundation explaining the emergence of relevant probability distributions, we may be able to make more accurate predictions.

Current efforts to model war dynamics

The current attempts to understand war dynamics are predominantly focusing on exploring the effect of capabilities on combat outcome. A second factor that frequently emerges in tandem with capabilities is their combat effectiveness referred in terms of firepower and specifically its accuracy. A most well-known and explored model is the Lanchester equation (both its linear and square versions), which has guided the modern military planning for decades (Bonder 2002, Kress 2012). Arguably, it is the Lanchester equation that gave the inertia to the practice of formally emphasizing either the preponderance of force or the effectiveness of troops' fire as determinants of combat outcome.

It may be useful to have a quick look at that model, especially since my own one would build on it. In the more popular Lanchester square law the impact of preponderance of force is emphasized. This is represented by a system of differential equations where *R* and *B* are the numbers of the belligerent forces (Red and Blue), while *r* and *b* are their respective fire effectiveness (Lepingwell 1987), interpreted as kill probabilities.

$$\frac{\partial \mathrm{B}}{\partial \mathrm{t}} = -\mathrm{rR}$$

$$\frac{\partial \mathrm{R}}{\partial \mathrm{t}} = -\mathrm{bB} \qquad (1)$$

We can see that the equations describe the attrition of belligerent forces over time as a function of the fire effectiveness of their opponent and the number of opponent's troops. Separating variables and integrating the above equations produces the illustrative square relationship, which emphasizes the impact of the level of forces on opponent's attrition rate.

$$\mathrm{rR}^2(0) = \mathrm{bB}^2(0) \tag{2}$$

To a large extent, it is this equation that added tension to the quantity over quality debate. As mentioned in many analyses of Lanchester equations, equation (2) indicates that increasing the number of initial forces *R(0)* and *B(0)* produces a higher effect on the attrition of the opposite forces than increasing the fire efficiency (*r, b*) of the forces (Lepingwell 1987). In fact, following (2) the incentive of the military planner would be to concentrate all possible forces against an opponent. This is because by increasing the speed of attrition of opponent's forces the military planner decreases its own losses, over time. In the following sections I will explore the speed of attrition variable more in detail.

The linear version of Lanchester's equation increases the importance of fire effectiveness. It does this by considering the attrition to be proportional to the degree of interaction of the two opponents' military capabilities over time.

$$\frac{\partial \mathrm{B}}{\partial \mathrm{t}} = -\mathrm{rRB} \tag{3}$$

$$\frac{\partial \mathrm{R}}{\partial \mathrm{t}} = -\mathrm{bBR}$$

This system of differential equations leads to the linear law version of the Lanchester state equation (4). It can be seen that both military strength and troops firing effectiveness have a comparable effect on the opponent's forces attrition.

$$rR(0) = bB(0) \tag{4}$$

Traditionally, the Lanchester linear law is believed to better represent ancient warfare characterized by individual duels. Indirect fire is another instance when its use is believed justified. In contrast, the square law is perceived to be more representative of the modern war where concentration of fire is possible (Dolansky 1964, Lepingwell 1987).

In this analysis I am going to adjust the Lanchester equations. Based on an analysis of combat interaction I will suggest a model that captures some essential dynamics not accounted for by Lanchester laws. The proposed adjustment responds to the criticism that Lanchester laws received over the years and attempts to address it.

One of the influential criticisms is reflected in Epstein (1985) work. He posed that while Lanchester equations have been useful in advancing our thinking about combat interaction, they suffer from a number of important shortcomings. Accordingly, Epstein indicates that none of the Lanchester models reflects an important feedback reaction in combat when one side's troops may withdraw to

avoid further attrition. In his words, this allows a side to trade space for time (1985, 6). It can be argued that this would allow a side to engage in combat at a different time and in more favorable conditions. In fact, Epstein concludes that in order for the Lanchester laws to have explanatory power there should be no movement of troops. That is, we are talking about stalemate combat interaction, of the type of trench warfare, which was so prevalent in WWI. Epstein brought the example of Iwo Jima fighting between Japanese and US troops (1985, 9) as an illustration.

Examining Lanchester equations Epstein suggested that each belligerent has to have the opportunity to withdraw its forces. This should happen in the case a belligerent noted that on the previous day it had achieved an attrition rate exceeding some set threshold value (1985, 16). According to Epstein such a model has a higher degree of fidelity and this indeed is more accurately capturing the combat dynamics.

But how is this relevant to the main question of this analysis? What Epstein invokes represents a much wider, even global underlying condition of any combat interaction. The trade-off between space and time that Epstein brings up represents an essential choice available to belligerents to either increase or reduce their exposure to opponent's lethal targeting. This exposure is one of the two key (and arguably innovative) concepts this analysis is going to introduce and develop. A good synthesis is that Lanchester models are less able to account for the role of strategy in battle outcomes.[19]

[19] This clarifying conclusion was suggested by a reviewer.

Withdrawing from combat and “trading space for time” is indeed very important, as it delays the attrition of one’s forces or it reduces the speed of that attrition. By offering more free-of-combat time to belligerents, this choice formally introduces a time lag into the exponential decay function of the Lanchester equations during which attrition is zero. This process is reflected in the decay graph by “injecting” into it some horizontal segments of zero attrition and creating a continuous step-shaped decay function. The “steps” operate by slowing down the attrition of less-exposed forces and create opportunities for them to “fight another day”. Thus, they increase the probability of reversing the troops attrition rate in particular belligerent’s favor. In other words, a lost battle does not imply losing the war. Or, the defeat in a skirmish does not necessarily translate into a lost battle.

Another influential research exploring ways to undermine the effect of general preponderance of force is reflected in the modeling of guerilla warfare by Deitchman (1962). He suggested that a side with an inferior number of troops can win against a military superior opponent, if it manages to concentrate superior forces or more effective firepower in distinct armed clashes (1962, 819). Deitchman argued that this is possible when the weaker side is ambushing the stronger side. Since the weaker side is hidden, it suffers an attrition in line with (4), because the stronger side should fire blindly in the area where the weaker side is located. In contrast, the attrition rate of the stronger side is governed by (2), because its troops are perfectly visible to the weaker ambushing side allowing the latter to conduct aimed firing. While Deitchman combines a system of equations using both the square and linear

law and invokes the importance of space, he nevertheless *also* models the combat dynamics based on capabilities and their fire effectiveness or accuracy. In fact, he explains how the weaker side can gain advantage by creating micro-cases of power preponderance. In political science literature this has been referred to as selection into battles that one believes can win. My proposed theory is different, as it suggests a weaker side can win even while not choosing to only fight the battles where it has power preponderance. It also reveals the conditions under which power preponderance could be a decisive factor on the battlefield.

To do this, I explore the spatial dimension of combat in a more fine-grained way. As reflected in the previous two examples, other similar attempts have been made. What is novel in my research, I would like to argue, is the extent of instrumentalizing the spatial factor. While political science research has also touched upon the relevance of spatial factor in warfare, besides military operations research, this has been done only superficially.

For instance, Biddle (2004) is probably one of the very few authors that explored specifically the role of spatial factor in explaining combat dynamics and how it can create variation in the effects of power preponderance. Biddle advances the idea of the “modern system” of combat, which implies cover, concealment, dispersion, small-unit independent maneuver, suppression, and combined arms

integration (2004, 35).[20] Except suppression, all the other listed actions explore the spatial dimension of combat.

However, while Biddle provided a solid qualitative description of the spatial dimension of combat and its importance, he did not fully explore this formally. He presented quite an intricate formal model of capability, which heavily explored firepower (capabilities) and their efficiency, much alike other described models did. His only variable exploring explicitly the spatial dimension of combat is f_e, which reflects a fraction of defender forces that are exposed, in his equation A.6 (2004, 212):

$$H = k_1(1 - f_e) \tag{5}$$

The *H* represent the number of attacker's troops that a dug-in and reinforced defender can halt. The constant k_1 is descriptively defined similarly to *H*. From that equation it is obvious that when fully exposed (f_e=1), the defender cannot halt any attacker, which is not in line with empirical data. In fact, under the condition of full exposure we would have a case when the classical Lanchester square law can accurately represent the combat interaction with casualties on both sides. Another limitation of Biddle's attempt to model exposure is that it has unilateral effect by affecting a segment of only the defender's forces. In reality, though, I argue and show that exposure is a global primitive that affects all sides involved in combat.

[20] Stephen Biddle is an atypical academic researcher, as he came into academia from the policy world, where he consulted the US military on defense planning and analysis. He has the comparative advantage of having been exposed to the most current practical issues and policy approaches in the US defense establishment.

It is worth mentioning that the formalization of combat interaction provided by Biddle seems to be more suitable to a concrete-purpose combat simulation rather than being a general description of combat dynamics. The attrition of troops in his model depends on technology, concentration of troops, reserves, velocity of troops movement and even assault frontage of the attacker. The model is too detailed to be able to claim describing and capturing modern combat dynamics in a general way. Many of these variables represent the same dynamic and can be reduced to describe a common process, thus decreasing this endogeneity in the model. This is a feasible thing to do because the approach triggers the question about why some tactics and parameters of combat are explored while others are ignored. It is creating an impression of arbitrary choice. Technology, for instance, serves as a multiplier for a few global variables including effectiveness of capabilities. In fact, Biddle choses to model policies and strategies that implement key combat dynamics and not the combat dynamics themselves, which is why his approach more resembles a simulation. Finally, I would like to argue that Biddle's qualitative description of combat dynamics focusing on reducing the exposure of troops, is theoretically more advanced and complete than other research examples listed earlier. However, his formal description of combat dynamics is less useful and explicit than those referred above. It does not manage to reflect well the conceptual depth of his qualitative explanation. In my following section I will attempt to address this shortcoming.

Explaining conflict outcomes through some spatial-related variables or concepts is not that infrequent in traditional political science. Stam (1996, 125) showed that terrain can have a significant impact on the military outcome. Supporting a similar logic in their empirical analysis of factors affecting the risk of civil war, Fearon and Laitin (2003) identified rough terrain as a condition favoring insurgency. Terrain has played an essential role in the arguments of influential geopolitical theorists (Mahan 1890, Mackinder 1904). Spatial considerations, through understanding the importance of geography, have significantly affected the military planning and foreign policy of many states. The United States has invested significant efforts and careful planning in setting up its military bases and defense cooperation abroad, for effective power projection. For instance, combat logistics and support – a key ingredient of any successful military campaign – depend on quick and unrestrained access to areas of expected combat engagement. As an example, Turkey's refusal to allow US troops to cross its territory during the 2003 invasion of Iraq[21] had affected the campaign-related costs for the United States.

Spatial considerations have greatly affected the choice of war and peace on the European continent. He defense-in-depth has been a historically-constant element of Russian security and defense policy. It tried to ensure it through conquering the neighboring countries (Trenin, 2001) or installing proxy governments, in order to operational them as buffer-sates (Mackintosh 1985,

[21] "Turkey Rejects US Troop Proposal," CNN.com, March 2, 2003, http://edition.cnn.com/2003/WORLD/meast/03/01/sprj.irq.main, accessed on 10 February 2019.

Ciziunas 2008). In economics, researchers working on conflict explored how spatial dimension can affect the propensity of war in an anarchic international system (Alesina and Spolaore, 2005).

It is my hope that I managed to convince the reader that spatial considerations are vital in affecting not only military and security policies of countries at the strategic level. I would like to argue that their political relevance and importance stems, especially, from their property of affecting the military outcomes, at both operational and tactical levels. Surprisingly, there has been little effort in political science research to provide a unified theoretical framework explaining the effects of space on a phenomenon so important for the discipline as the crisis escalation is. While there is sufficient empirical work suggesting the importance of space in affecting war costs and war outcome, the existing theoretical efforts to explain the mechanism of this effect are still inconclusive.

Exposure – the silent factor

The degree of exposure, as a global primitive of the combat dynamics, has a "silent" effect on combat outcome. While it has practically complete control over how power preponderance can influence war outcomes, exposure has always been the "invisible hand" of war. I shall illustrate its effect on war outcome further in this section.

I will start by building the logical context for my analysis by explaining the microdynamics of conflict, which should allow me to define and better present the exposure primitive. For this purpose we will have to zoom in on combat interaction at a very disaggregated level. I will provide the context necessary to understand combat dynamics by referring to some existing work, to give the well-deserved credit. While I conceptualized similar ideas myself, I later found they were reflected in one form or another in the works listed below.

According to Ancker (1995, 318) the basic building block of combat interaction is the firefight. He poses that everything else is secondary and only supports the effort to win the war. To better place it into the war dynamics context, he suggested the following logical hierarchy: firefight, engagement, battle, campaign, and war. This is less relevant for this work's larger argument, but it helps gain a more accurate perspective.

Ancker provides a logically rigorous description of the combat process through his two axioms "for a theory of combat". His first axiom introduces what he labels "the fundamental property of the system" (1995, 321). It states that a firefight is just a union of all sets of pairs containing the firer and its accessible target and which have non-empty intersections. In his conceptualization, the firer employs lethal means against at least one target on the opposite side. The target is accessible and vulnerable to the lethal impact of at least one opponent firer. The accessibility qualifier simply indicates those targets that are both identified and are in the range of lethal means of the firer (figure 2.1). Being just in the range of the

lethal strike does not necessarily make a side targeted if the latter is not identified. This has a critical importance to my argument because identification is dependent on the spatial phenomena that makes evasion possible and thus is having a strong impact on exposure. It is worth mentioning that a target can be at the same time a firer and a target.

Ancker's second axiom describes the dynamic process of the firer-target engagement. It states that a firefight is nothing but a "terminating stochastic target attrition process" that is continuous in its nature. This echoes well with Hirshleifer (2000) statement that was quoted at the beginning of the analysis, claiming all combat at the point of contact with the enemy is reduced to troops attrition.

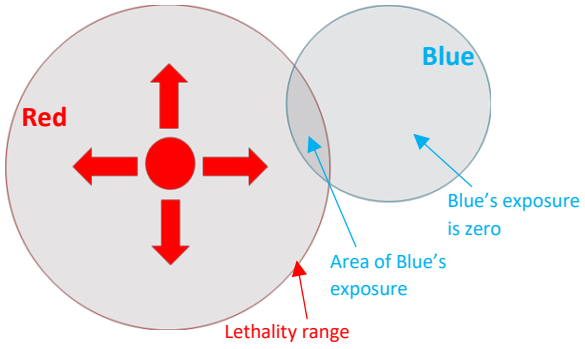


Fig. 2.1: Firer's lethality and target's vulnerability and exposure

It is important to conceptualize correctly that attrition occurs only at the intersection of the two sets: the one set limiting the firer lethality range and the other one delineating the target's physical presence and vulnerability.

As a matter of clarification, if the target is outside the firer lethality range it will not suffer any losses. If the target is inside of the firer lethality range but not perceived by the firer to be there, the target will not be acquired and will avoid losses, again. However, if the target is *both* in the firer lethality range and is exposed to the attack either because is visually identified or simply believed by the firer to be in the area where strikes can be delivered, the target will mostly likely suffer losses.

The firer entity has lethality as an active property while the target entity has the passive property of vulnerability (Ancker 1995, 322). As derived from Ancker's two axioms, vulnerability depends on spatial factors. It is enacted by the target stepping into the area of firer lethality range and by being observed. Both location and visibility are determined by space parameters. This is showed by Biddle (2004, 53), who indicates that direct-fire weapons operate on loosely flat trajectories, which makes it possible to explore the Earth's surface to achieve cover and concealment. Simpler, this means that a bullet flies straight and it cannot hit an infantryman laying down below that trajectory. This creates opportunities to reduce the exposure of the target. Any kind of obstacle in the bullet's way (a tree, a hill, a building, or armor) then is a spatial element allowing to decrease the target's exposure to lethal fire and to increase its evasion.

This elaborate conflict microdynamics description was necessary because it allows to better visualize how spatial dimension matters and how it allows for the emergence of exposure (in terms of vulnerability) property. Biddle (2004, 54) convincingly reveals that even against technologically advanced weapons system,

the lethality can be decreased by exploring spatial parameters to reduce exposure in comparison with the targets that are massed in the open. Reduction of exposure can be achieved also by using technical countermeasures, which suggests that technology is a simple force-multiplier that can amplify either lethality or exposure.

Some illustrative examples describing active decrease of troops exposure from high-tech reconnaissance efforts includes operation ANACONDA in Afghanistan in 2002. In that instance the Taliban managed to hide their positions very well and avoid detection and destruction before the direct troops' engagement (Biddle 2002, 28). A less known example is the very successful decrease of its forces' exposure by Yugoslav military during the 1999 US-led NATO airstrikes, employing decoys and engaged battlefield disinformation.[22] An accurate summary of the conflict microdynamics description would be that strategy is then about securing advantageous firefights and avoiding those at which one is at disadvantage. The advantage in firefights is achieved by increasing the opponent's exposure to fire through limiting its degrees of freedom and by decreasing one's own exposure to fire. Or, as Hirshleifer (2000, 775) suggested, one side should arrange its forces in space and time in such a way that, should a battle occur, the attrition rates should favor its forces.

Up until now I claimed that space-induced properties have an essential effect on the ability of belligerents to inflict lethal damage against each other, by either

[22] "The World; Bombs are Smart. People are Smarter," The New York Times, July 4, 1999, https://www.nytimes.com/1999/07/04/weekinreview/the-world-bombs-are-smart-people-are-smarter.html, accessed on 12 January 2019.

increasing it or decreasing. I also provided a qualitative clarification of this claim. To make this more obvious a more formal description and analysis are necessary.

Given that the application of firepower, either generated by one firer or a concentration of troops, is spatial, fire is going to be directed towards specific territorial area. In fact, we can claim that firepower application has spatial projection constrains being largely unidirectional. At the most basic level, when one individual is facing two or more opponents in a physical contest, given similar skills, she can address them effectively only one at a time. A bullet can only be fired in one direction. An artillery projectile can be launched similarly in one direction. Even a salvo from a multiple-launch rocket system, is still unidirectional.

I will represent this logic in the illustration from figure 2.2. Imagine that it represents a conceptual battlefield, where Red's forces are occupying a single battlespace, S_r, in which they are visible and are having a very high degree of exposure, $\mathcal{E}_r$. The Blue's forces are dispersed among multiple battlespaces, from S_{b1} to S_{bn}.

The Blue's forces then have a varying degree of exposure $\mathcal{E}_{b,}$ depending on which space S_b they are located in. The reference to the (battle) "space" should not be taken literally, as this is an abstraction. For instance, some of Blue's troops may be established in defense positions – one of possible battlespaces - well known to Red and thus having a higher exposure to lethal fire. Defense fortifications would partially decrease that exposure. Armored vehicles and troops they protect have a more reduced exposure against munitions that cannot penetrate the armor.

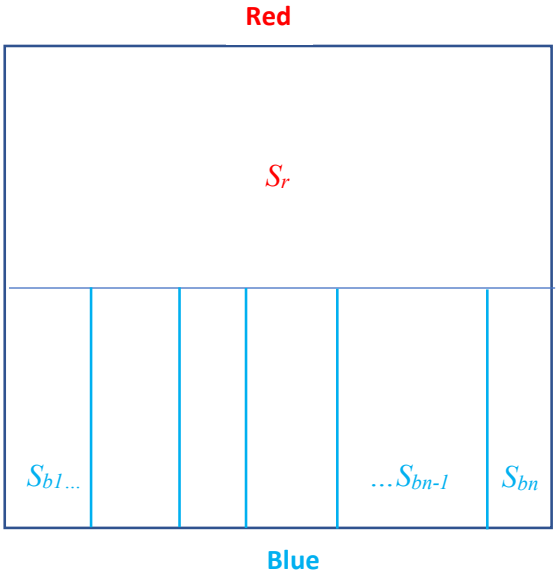


Fig. 2.2: Conceptual distribution of Red and Blue forces across battlespaces S_r and S_b

Those Blue troops that are dispersed across locations (or battlespaces) unknown to Red have a minimum degree of exposure. Some other troops can be located in inhabited areas – another possible battlespace.[23] Or, they may be moving to attack the rear and logistical lines of Red forces, being situated in yet one more type of conceptual battlespaces.

Let us consider the hypothetical situation when Red does not know in which of the *n* battlespaces the Blue forces are located. Therefore, the chance to hit Blue forces would be *1/n*, if only one battlespace can be hit at a time and Blue is known to occupy just one of the spaces. Or, the chance of hit would be *m/n*, if Blue's forces are scattered across *m* battlespaces and more than one can be targeted by Red. It

[23] International Humanitarian Law has significant restraints about the use of indiscriminate firepower (such as artillery or air strikes) in inhabited area, which makes these areas a battlespace of conditionally limited combat exposure for a defender, unless the attacker wants to violate IHL and pay the political costs of targeting inhabited areas indiscriminately.

shall be considered that hitting more than one battlespace increases the costs for Red, as it has to disperse its troops to address each targeted battlespace by a separate sub-set of its forces. Targeting specific battlespaces can be also costly if these happen to be inhabited areas. This could trigger potential audience costs due to the destruction of civilian infrastructure and casualties among non-combatants, violating International Humanitarian Law provisions. Therefore, it can matter if Red has to target just one battlespace or a few.

As earlier suggested in figure 2.1 Blue's attrition and thus exposure to Red's lethal fire depends on its ability to avoid identification by Red. Therefore, Blue's degree of exposure becomes basically dependent upon Red's knowledge about Blue's location, which reflects well Anker (1995). The higher the uncertainty of Red about Blue's location, the lower the exposure of Blue's forces becomes. This, in turn, would diminish the attrition of Blue's forces and allow it to continue fighting for longer periods.

One way to formally represent Red's (lack of) knowledge about Blue's location is through information entropy, developed by Shannon (1948). In its original sense, informational entropy is explaining how much uncertainty is there about the system of interest, or how much information is needed to know the states (configurations) of a complete system of events (Reza 1961, 77-78). Shannon suggested that:

$$H = -\sum_{i=1}^{n} p_i \log p_i \qquad (6)$$

in which p_i is the probability that an event occurs, and $0 \leq p_i \leq 1$. The logarithm with binary base is used, which is also convenient for the issue addressed in this section, given it will explore a binary choice.

Using the above formula for information entropy I could consequently define more formally the concept of combat or lethal exposure that I introduced in this section. Re-writing the above information entropy formula to represent Red's uncertainty about Blue's location over battle spaces (S_{bi}), as illustrated in figure 2.2, gives us:

$$H_r = -\sum_{i=1}^{n} S_{bi} \log S_{bi} \tag{7}$$

The formula above also is able to provide insights into the exposure rates of either Blue ($\mathcal{E}_b$) or Red ($\mathcal{E}_r$) forces, depending on opponent's knowledge of its troops distribution across battlespace (S_b, S_r) where its troops are located.

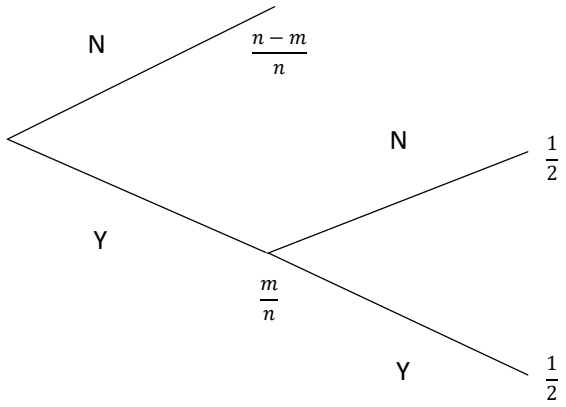


Fig. 2.3: Probability scheme of belligerent's uncertainty about its opponent's location

To be able to evaluate and compare the exposures of the two sides and accurately appreciate the role of exposure on the process of attrition, I would like to present in figure 2.3 what is called a probability scheme (Reza 1961). It illustrates the *uncertainty about the outcomes* of the expected combat interactions between Blue and Red forces.

To clarify the intuition revealed by figure 2.3, imagine that Red does not know the location of the Blue forces, but knows they are distributed across *n* possible battlespaces, from S_{b1} to S_{bn}, as implied in figure 2.2. Based on the current literature on conflict, I suggest that in order to assess its chances of victory over Blue, Red needs to know a few things. This includes the location of Blue as well as whether Blue will engage Red directly or through indirect tactics. The second part is relevant, since the amount of lethal exposure is different when comparing the direct conventional combat and the indirect, insurgency-based interactions.

There is a different way to conceptualize this. The uncertainty about Blue's location across the set of battlespaces S_{bi} would also include uncertainty about the unknown nature of Blue's forces. The assumption, which is reflective of practical experience, is that Red generally knows the overall size of Blue forces but he does not know *the partition* of Blue forces across specific battlespaces. This is why in fig. 2.3, at the second node of the probability tree, Red tries to gauge the type and size of the force it may face because that would determine the tactics of Blue against Red.

Therefore, the model assumes that in order to form an accurate perception about its probability of winning, the militarily-stronger Red would need to know where Blue forces are located and what tactics they are going to choose to fight Red - conventional or insurgency tactics. In line with Shannon's information entropy logic, the uncertainty is reduced by asking a number of "yes" or "no" questions. As a first step (1) Red would examine each individual battlespace *n* and enquire whether Blue forces are located there or not. This may be achieved through intelligence gathering, for instance. Figure 2.3 denotes a more general case, when Blue's forces are scattered across *m* of *n* battlespaces. By enquiring over all *n* possibilities Red would estimate the involved uncertainty by identifying a set of *m* battlespaces where Blue is located. Thus, Red will also clarify the spaces where Blue forces are not located in *(n-m)/n*. Following this, as a second step (2) Red would examine each battlespace where Blue forces are likely to be and enquire whether Blue will face Red in conventional or unconventional combat. For illustrative purpose, and simplicity, let us assume that Red is not able to get more clarity into this issue and there is an equal probability that Blue will fight him either conventionally or unconventionally, $p=½$.[24] Or, he will be faced half of the time by an opponent using conventional tactics and half of the time by one employing insurgent tactics.

[24] This probability can be interpreted in a few feasible ways. For instance, it can represent the maximum uncertainty as to how the opponent is going to fight the Red or about his particular force partition across battlespaces. If we are referring to an insurgent, we can view this as the fact that the insurgent will attack the Red forces in a conventional fashion, but at his own choosing, attempting to catch the Red in tactically inconvenient situations or amassing local superiority of forces. For instance, Biddle (2002) suggested that the Taliban would occasionally explore conventional tactics to attack US or Afghan Government forces. The choice between conventional or insurgent sometimes is made based on the tactical and operational objectives, as well as the local concentration of forces.

This second enquiry is important because, as indicated in Deitchman (1962), one of the possible ways to achieve maximum attrition of the opponent's forces is by concentrating superior number of troops against them, at the point of combat contact. If Red is able to get Blue to fight conventionally, it will have higher chances to bring superior number of troops against Blue. It will thus be able to better explore its military advantage, given its assumed general preponderance of forces. In the case when Blue is likely to fight unconventionally, for instance through the use of low-exposure tactics such as hit-and-run, it will do so at the time and place of its choosing frequently ambushing Red's forces. Contrary to Deitchman's model, my model's findings indicate that the ambusher does not necessarily need superiority in forces to win against the ambushed. It just need an advantage in exposure rate.

Applying equation (7) we can calculate Red's uncertainty about the outcome of a confrontation with Blue's forces. This is achieved by understanding Blue's exposure, its rate of attrition and consequently the probability of Red's victory. To achieve this I propose the following application:

$$H_r = -(\tfrac{1}{2}\,\tfrac{m}{n}\,\log(\tfrac{1}{2}\,\tfrac{m}{n}) + \tfrac{1}{2}\,\tfrac{m}{n}\,\log(\tfrac{1}{2}\,\tfrac{m}{n}) + \tfrac{n-m}{n}\,\log\tfrac{n-m}{n}) \tag{8}$$

$$H_r = -(\tfrac{m}{2n}\,\log\tfrac{m}{2n} + \tfrac{m}{2n}\,\log\tfrac{m}{2n} + \tfrac{n-m}{n}\,\log\tfrac{n-m}{n}) \tag{9}$$

Simplifying we receive:

$$H_r = -\left(\frac{m}{n}\log\frac{m}{2n} + \frac{n-m}{n}\log\frac{n-m}{n}\right) = \frac{1}{n}\left(n\log\frac{n}{n-m} - m\log\frac{m}{n-m} + m\right) \quad (10)$$

To be able to make sense of this abstract image of entropy and consequently exposure, I am replacing the variables of interest with concrete values. Let us take the example when *n*=16 and *m*=4. Then we could re-write (10) as follows:

$$H_r = \frac{1}{16}\left(16\log\frac{16}{16-4} - 4\log\frac{4}{16-4} + 4\right) \approx 1.07 \quad (11)$$

For comparison, if we take the example when Blue's forces are only located in one of the battlespaces S_{bn}, we have the following result:

$$H_r = -(-\frac{1}{n}\log 2n + \frac{n-1}{n}\log\frac{n-1}{n}) \quad (12)$$

Substituting the selected measures, we receive the following:

$$H_r = -\left(-\frac{1}{16}\log(2*16) + \frac{16-1}{16}\log\frac{16-1}{16}\right) \approx 0.404 \quad (13)$$

Finally, for a better clarification of the combat entropy logic, let us examine the uncertainty of Blue about Red's troops location, given figure 2.2. Since Red is known to have concentrated its troops in the battlespace S_r, using (7) for Blue's uncertainty we get:

$$H_b = -\left(\frac{1}{1}\log\frac{1}{1}\right) \to 0 \quad (14)$$

While initially the results from (11) and (13) may seem puzzling, there is an observable logic in them. When Blue's forces are scattered wider, across more battlespaces, the uncertainty of Red about Blue's forces location is higher. In other words, Red needs more information to learn completely the Blue's forces location and their possible combat tactics. When Blue is more concentrated in the pattern of its force deployment, Red needs less information to identify the bulk of Blue troops, its expected combat tactics, and thus has a lower level of uncertainty. To portray this even more differently, consider that when Blue forces are scattered across multiple locations, targeting one location would only possibly destroy a small part of the Blue's forces. However, if Blue is only concentrated in one location, Red's ability to identify and target that location will cause a much more significant damage, thus revealing a higher degree of Blue's troops exposure.

Based on the intuition provided by the information entropy logic, we can now get a better understanding of the combat exposure concept that this section introduces. In other words, we can evaluate Blue's (or Red's) exposure rates. The exposure measure, as further illustrated, will be of an ordinal nature though, as it will allow us to only make comparisons among different combat scenarios, given the combat-related entropy the sides have in relation to each other.

I am defining Blue's exposure as the inverse function of the degree of uncertainty that Red has about Blue's position. If Red knows Blue's position, this increases Blue's exposure to enemy fire and accelerates the attrition of its troops.

Presenting exposure of one side as being inversely proportional to its opponent's uncertainty about the location of its forces, we will have

$$f(\mathcal{E}_b) = f(H_r)^{-1} \quad (15)$$

For instance, given the certainty of Blue forces about the location of Red, $H_b \rightarrow 0$, the exposure of the Red forces $\mathcal{E}_r$ will be at maximum value. In other cases, in which we measured Red's entropy about Blue's location $H_r \approx 0.404$, Blue's exposure will be $\mathcal{E}_b \approx 2.47$; and when the value of $H_r \approx 1.07$ we will have $\mathcal{E}_b \approx 0.93$. While the developed logic is not sufficiently mature to allow yet making exact sense of the numerical value of the figures, they are still offering a number of practical insights. They are useful to demonstrate that as a side's forces are distributed across multiple combat spaces, the opponent's uncertainty about the location of these forces is increasing. This also is decreasing the exposure of that side's forces to lethal engagement from the opponent. The figures show that the higher the uncertainty of Red about Blue location is, the lower is the Blue exposure to Red lethal fire. By applying the information entropy logic, I show that attrition rate of troops in combat is critically dependent on these troops' exposure. In fact, very low exposure could totally cancel the impact of the opponent's force preponderance and its fire accuracy.

How do we translate this abstract concept into physical realm of combat? A decreased exposure can have multiple physical manifestations, including engaging in combat at a distance, at which the opponent cannot reach one's forces with lethal

fire, maneuvering to avoid fortified positions of the opponent, engaging the opponent from an ambush, fortifying its own positions including through the use of armor that the opponent's weapons have difficulty to penetrate, etc. For instance, the German forces bypassing the French Maginot Line through the less fortified Low Countries was aimed at reducing the exposure of the German forces to French lethal fire. Luttwak (2001, 5) gave an illuminating explanation of the exposure logic, which he presents in the framework of generating surprise. He suggested that maneuver is not a mere movement, but an attempt to "circumvent the greater strengths of the enemy and to exploit his weaknesses". Or, in the interpretation of this analysis the goal is to adjust the combat interaction conditions such that one's own exposure is minimized, while the opponent's exposure to lethal fire is maximized. In the presentation of this analysis exposure is a global primitive of combat dynamics. It subsumes all the tactics described as exposure's physical manifestations at the beginning of this paragraph and includes some others. The novel part of my analysis is that it separates exposure as a critical and perhaps the most important primitive of combat dynamics. In contrast, other research examined some elements of exposure, while ignoring others and fail to emphasize its primacy. Or, they were not able to separate exposure as a global primitive of combat.

To summarize, exposure is the logical descriptor used to explain the degree to which a conflicting side's troops are subject to lethal fire of the competitor, resulting in a varying level of attrition of its troops greater than zero. A higher exposure will result in higher attrition rate of its troops, and vice versa. Using figure

2.1 exposure is determined by both the physical presence of a combatant in the opponent's range of fire *and* the knowledge of the opponent about that combatant's location in the targeting range.

The main goal of this section was to indicate that apart from the existing factors used to measure the level of attrition – preponderance of forces and their targeting accuracy – a very significant impact on attrition is carried by the degree of troops exposure. In fact, when the degree of troops exposure is approaching zero and the opponent does not know what area to target with lethal fire, the attrition will be close to zero. It will have this value regardless of power preponderance and superiority in fire accuracy. As exposure increases the effects of the quantity of troops and the targeting accuracy of the opponent increases too, thus allowing for their effects to materialize, if they exist. Therefore, I would like to claim that in order to accurately represent modern combat one needs necessarily to consider what I labeled as the troops' exposure. This section's analysis allows us to claim that targeting effectiveness of one side completely depends upon the value of exposure of the opponent.

As a conclusion, I would like to claim that what Lanchester used to refer in his equations as combat effectiveness is, in fact, a proxy variable that partially reflects the effects of troops exposure. The troops' fire effectiveness will only materialize in close combat, where exposure will naturally be increased. Thus, when two soldiers see each other and are in the range of fire of each other's weapons their exposure is at its maximum value. In the next section I intend to review the

existing combat analysis equations by adding to them the variable of exposure and analyzing the resulting effect on combat attrition.

Proposing a general model of combat

Modern combat analysis generally focusses on identifying and measuring the effect of various factors on the degree of attrition of forces engaged in combat. The logic is quite simple. As an aggressor's army marches towards the capital of the opponent's country to capture it and declare victory, or as it moves towards taking control over a strategic segment of territory (offering control over resources or defense advantages) the only obstacle in its way is the army of the defending country. Thus, inflicting damage to that force (through combat attrition) allows to remove this obstacle. Another solution would involve the defender surrendering without fighting. But this would only be possible, rationally, if the defender reaches the conclusion that as a result of expected combat its troops will be destroyed first. So, the expectation of one's forces quicker combat attrition, based on similar understanding of combat dynamics of the belligerents or common knowledge, could have an effect similar to military victory. It is about this situation that Sun Tzu is believed to have said that "the greatest victory is that which requires no battle". Empirically, the described logic has had a significant effect on military commanders view on combat (Malkasian, 2002).

I will base my model on the Lachester equations logic due to a few reasons. First, Lanchester equations continue to be heavily used in military operations research and planning. Then, even the staunchest critics of the model recognize that it captures well the combat dynamics at the point of combat contact (Epstein 1985, 9) in trench type of warfare. Finally, it generally represents combat attrition rather well, as a continuous degradation of forces process. It is also worth mentioning that Lanchester equations represent a typical exponential decay model, which have been widely and successfully used in evolutionary biology to model competition among species. Therefore, it is a logically and empirically good starting point.

However, based on the results shown in the previous section it is clear that exposure may have a significant impact on both the degree and the speed of attrition in combat. Therefore, I will propose the following system of coupled differential equations to describe combat between two opponents, Blue and Red:

$$\frac{\partial \mathrm{B}}{\partial \mathrm{t}} = -\mathcal{E}_{\mathrm{b}}\mathrm{B} - \mathrm{rRB}$$

$$\frac{\partial \mathrm{R}}{\partial \mathrm{t}} = -\mathcal{E}_{\mathrm{r}}\mathrm{R} - \mathrm{bBR} \tag{16}$$

As it can be seen, the proposed model also includes the degradation of Blue forces by Red through attrition, dependent on the exposure coefficient $\mathcal{E} \geq 0$. Also, it makes attrition of sides dependent on the level of interaction of their troops with

the opponent's troops. This means, the number of killed troops depends both on what number of forces engage in combat, but also on the size of the engaged opponent, which echoes Sullivan (2007). This, as suggested by biological modeling of species competition, is empirically a more valid assumption. The other two constants of the system are *r*, *b* ≥ *0*, representing the effectiveness or accuracy of fire of *R* and respectively *B* forces, which are the independent variables, *R*, *B* ≥ 0.

The dynamic system above consists of a system of ordinary differential equations. Because they are nonlinear, they are either difficult to solve (Rubinow 2003, 24) or not always even possible to solve analytically (Strogatz 2015, 8). However, there have been developed effective approximation methods that can give approximate solutions, which are useful for the understanding of the system. I will use the linearization technique with the aim to solve the system analytically (Strogatz 2015, 151; Noonburg 2014, 161) and consequently observe its behavior in a phase plane. For that purpose, I will transform the system of equations from (16) into a linear system of the form $\vec{\mathbf{x}}' = \mathbf{A}\vec{\mathbf{x}}$. For a more familiar mathematical representation I will replace *B*=*x* and *R*=*y*:

$$\frac{\partial \mathrm{x}}{\partial \mathrm{t}} = -\mathfrak{C}_{\mathrm{b}}\mathrm{x} - \mathrm{rxy}$$

$$\frac{\partial \mathrm{y}}{\partial \mathrm{t}} = -\mathfrak{C}_{\mathrm{r}}\mathrm{y} - \mathrm{bxy} \qquad (17)$$

As a first step, I will examine the fixed points of the system, by setting the right-hand side null:

$$\mathrm{x}(-\mathfrak{C}_{\mathrm{b}} - \mathrm{ry}) = 0$$

$$\mathrm{y}(-\mathfrak{C}_{\mathrm{r}} - \mathrm{bx}) = 0 \tag{18}$$

The solutions for a trivial fixed-point are *x*=0 and *y*=0. The less informative fixed-point values are representing another equilibrium state and are $(-\frac{\mathfrak{C}_{\mathrm{r}}}{b}, -\frac{\mathfrak{C}_{\mathrm{b}}}{r})$, which make less analytical sense in regard to the examined model and thus are not explored.

I will then compute the Jacobian of (17) in order to classify the fixed point through linearization:

$$\mathrm{A} = \begin{pmatrix} \frac{\partial \dot{\mathrm{x}}}{\partial \mathrm{x}} & \frac{\partial \dot{\mathrm{x}}}{\partial y} \\ \frac{\partial \dot{\mathrm{y}}}{\partial \mathrm{x}} & \frac{\partial \dot{\mathrm{y}}}{\partial y} \end{pmatrix} = \begin{pmatrix} -\mathfrak{C}_{\mathrm{b}} - \mathrm{ry} & -\mathrm{rx} \\ -\mathrm{by} & -\mathfrak{C}_{\mathrm{r}} - \mathrm{bx} \end{pmatrix} \tag{19}$$

Provided the fixed point (*x, y*) = (0, 0), the Jacobian matrix at these coordinates is:

$$\mathrm{A} = \begin{pmatrix} -\mathfrak{C}_{\mathrm{b}} & 0 \\ 0 & -\mathfrak{C}_{\mathrm{r}} \end{pmatrix} \tag{20}$$

I then identify the eigenvalues by finding first the trace of the matrix *A*, $\tau(A) = -\mathcal{E}_b - \mathcal{E}_r$ and the determinant $\Delta(A) = \mathcal{E}_b\mathcal{E}_r$. The eigenvalues can be determined according to the following formula (Strogatz 2015, 132):

$$\lambda = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2} \tag{21}$$

$$\lambda = \frac{-\mathcal{E}_b - \mathcal{E}_r \pm \sqrt{(-\mathcal{E}_b - \mathcal{E}_r)^2 - 4(\mathcal{E}_b \mathcal{E}_r)}}{2} \tag{22}$$

This suggests that $\lambda_1 = -\mathcal{E}_b$ and $\lambda_2 = -\mathcal{E}_r$, which makes the fixed point at (x, y) = (0, 0) a stable node.

To find the eigenvectors, corresponding to the eigenvalues $\lambda_1 = -\mathcal{E}_b$ and $\lambda_2 = -\mathcal{E}_r$ we have to set the following equation equal to zero:

$$(\mathbf{A} - \lambda \mathbf{I})\vec{\mathbf{v}} = \mathbf{0} \tag{23}$$

which for λ_1 results in

$$(\mathbf{A} - (-\mathcal{E}_b)\mathbf{I})\vec{\mathbf{v}} = \begin{pmatrix} -\mathcal{E}_b - (-\mathcal{E}_b) & 0 \\ 0 & -\mathcal{E}_r - (-\mathcal{E}_b) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & -\mathcal{E}_r + \mathcal{E}_b \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ v_2(\mathcal{E}_b - \mathcal{E}_r) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{24}$$

and for λ_2 we will have respectively

$$(\mathbf{A} - (-\mathcal{E}_r)\mathbf{I})\vec{\mathbf{v}} = \begin{pmatrix} -\mathcal{E}_b - (-\mathcal{E}_r) & 0 \\ 0 & -\mathcal{E}_r - (-\mathcal{E}_r) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} =$$

$$= \begin{pmatrix} \mathfrak{E}_r - \mathfrak{E}_b & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix} = \begin{pmatrix} \upsilon_1(\mathfrak{E}_r - \mathfrak{E}_b) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (25)$$

receiving the eigenpairs $\left(-\mathfrak{E}_b, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$, and $\left(-\mathfrak{E}_r, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$.

Following that we can write the general solutions for the examined system, as:

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{-\mathfrak{E}_b t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{-\mathfrak{E}_r t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (26)$$

This then results in the following solutions for each Red and Blue forces:

$$x(t) = c_1 e^{-\mathfrak{E}_b t};\ \ y(t) = c_2 e^{-\mathfrak{E}_r t} \quad (27)$$

In that equation c_1 and c_2 are constants of integration, reflecting the initial conditions of the system such as the initial number of troops. A complete phase portrait of this system[25] (17) is presented in figure 2.4a, though we are interested in the first quadrant, as shown in figure 2.4b.

It is informative though to have a look of the phase portrait of the whole system, given that its eigenvalues generate a more complex picture, with a second fixed point at $(-\frac{\mathfrak{E}_r}{b}, -\frac{\mathfrak{E}_b}{r})$, which is visible in figure 2.4a. Figures 2.4b and 2.5a represent the same system dynamics, with the only difference that 2.5a displays a zoomed-in perspective, so that we can better see the effect of eigenvalues on the degradation of the combat resources (troops) of the belligerents.

[25] The phase portrait was generated using Wolfram Alpha online applet, using $\mathcal{E}_b$=.2, $\mathcal{E}_r$=.6, while b=r=.4.

That phase portrait illustrates the intuition that depending on the size of the eigenvalues, the degradation of the sides' forces can be slower or quicker over the time. In figure 2.5a we can see that Blue's forces degrade slower (near x axis) than Red's forces (near y axis), as the direction field reveals.

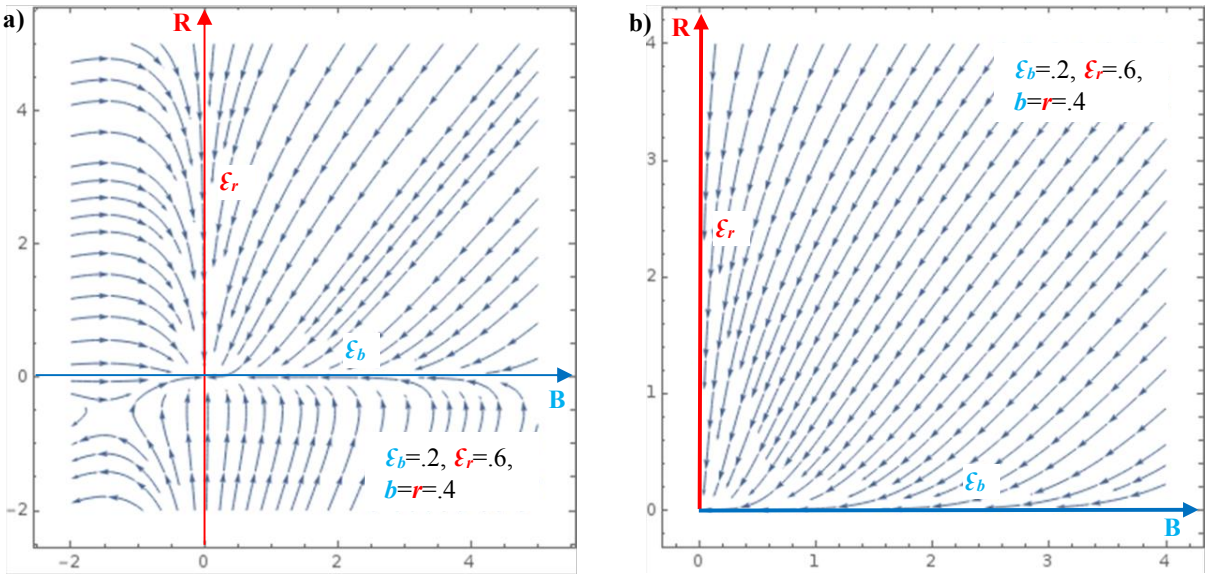


Fig. 2.4: Phase portrait of the system of equations (17): $\dot{x}$=-0.2x-0.4xy; $\dot{y}$=-0.6y-0.4xy

This is due to the fact that close to the node trajectories are tangential to the slow eigendirection, which is the eigendirection with the smallest eigenvalue (Strogatz 2015, 157). In dynamic terms, the reduced "speed" of the lower eigenvalue allows it to approach the *x* axis well before the origin. To better understand this dynamic interaction and the effect of eigenvalues (thus of the exposure rates) on the speed of side's troops attrition, let us take a closer examination of the figure 2.5a.

The black diagonal line represents the coordinates where the number of Blue and Red forces are equal. It can be seen that the lower exposure of Blue comparing to Red ($\mathcal{E}_b = 0.2 < \mathcal{E}_r = 0.6$), while keeping combat effectiveness constant ($b = r = 0.4$), results in a considerably quicker attrition of Red troops.

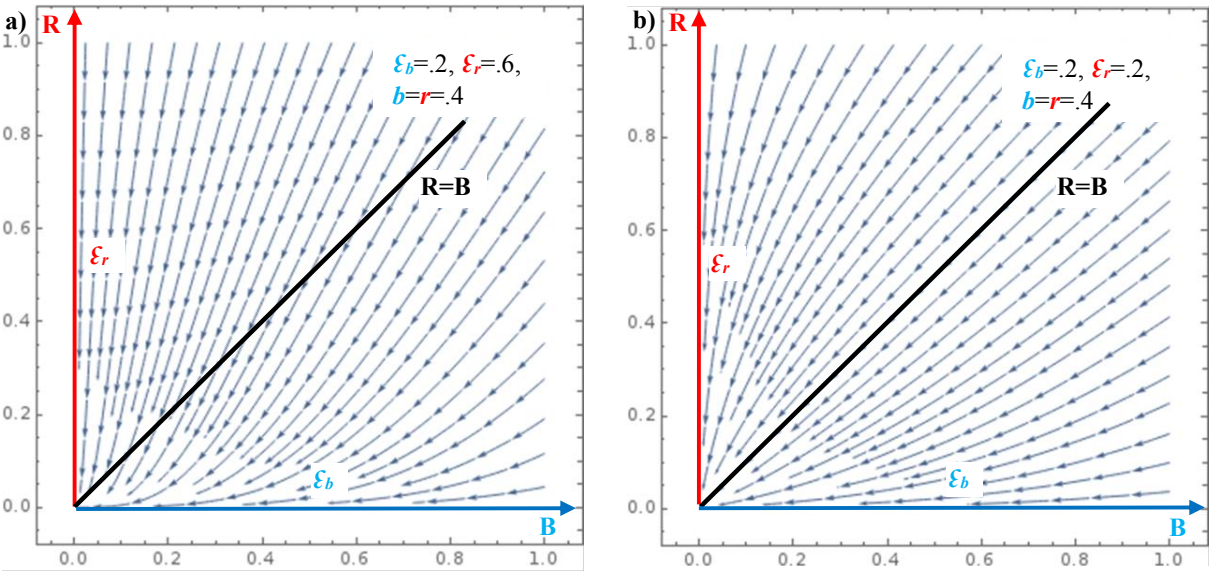


Fig. 2.5: Phase portrait of the system of equations (17): I quadrant with different $\mathcal{E}_i$[26]

For instance, even if we start with 0.2 units of Blue troops and 1 unit of Red troops (figure 2.5a), the Blue forces are in fact likely to stand on the battlefield after all Red forces are destroyed. This is due to the effect of eigenvalues, which is more pronounced closer to the fixed point. In contrast, the trajectories of the direction field in the upper-right corner of the quadrant look parallel. This happens because they reflect the effect of the nonlinear interaction terms (rxy and bxy) of the equation (17)

[26] The two graphs are described by the following dynamic systems: $\dot{x}$=-0.2x-0.4xy; $\dot{y}$=-0.6y-0.4xy (fig.5a) and $\dot{x}$=-0.2x-0.4xy; $\dot{y}$=-0.2y-0.4xy (fig.2.5b).

on the dynamical system. There is an empirical logic in that too. This is because in direct close combat, when troops of the opposing sides observe each other, the value of exposure increases naturally due to less opportunities to reduce it.

The impact of troops' fire effectiveness may also frequently become non-significant. This can especially happen during massive fire exchanges, like in artillery strike or air-bombing. However, even in these cases is possible to explore the terrain irregularities as well as dispersal of troops aimed to evade some of the enemy fire, thus keeping exposure well below a maximal value (Biddle 2004).

To better clarify that dynamic system and the effect of eigenvalues, let us modify the values of $\mathcal{E}_i$.[27] In figure 2.5b, the exposure of Red and Blue forces, and thus the eigenvalues of the system are equal $(\mathcal{E}_b = \mathcal{E}_r = 0.2)$. We witness a completely different picture comparing to figure 2.5a. The eigendirection produces identical attrition for both sides. In fact, it is obvious that at the equal level of exposure the two forces are degrading symmetrically.

Let us examine two extreme cases, as these are able to better reveal the effect of the variables of interest. In figure 2.6a and 2.6b below we have minimal exposure for Blue forces and maximum exposure for Red forces $(\mathcal{E}_b = 0.1 < \mathcal{E}_r = 0.9)$ and $(\mathcal{E}_b = 0.01 < \mathcal{E}_r = 0.9)$ respectively.[28]

It is clear from the most extreme case in figure 2.6b, that even if Blue starts with 0.05 units of forces, while Red starts with 1 unit of forces (a 20:1 advantage in

[27] To remind the reader, the exposure, $\mathcal{E}_i$, also determines the eigenvalues of the system $(-\mathcal{E}_b, -\mathcal{E}_r)$.
[28] This corresponds to the following dynamic systems: $\dot{x}$=-0.1x-0.4xy; $\dot{y}$=-0.9y-0.4xy and $\dot{x}$=-0.01x-0.4xy; $\dot{y}$=-0.9y-0.4xy, respectively.

numbers for Red), the Blue loses almost none of its forces, while the Red forces are annihilated. A similar effect is noticed in figure 2.6a, though the loses of Blue forces are higher in this case.

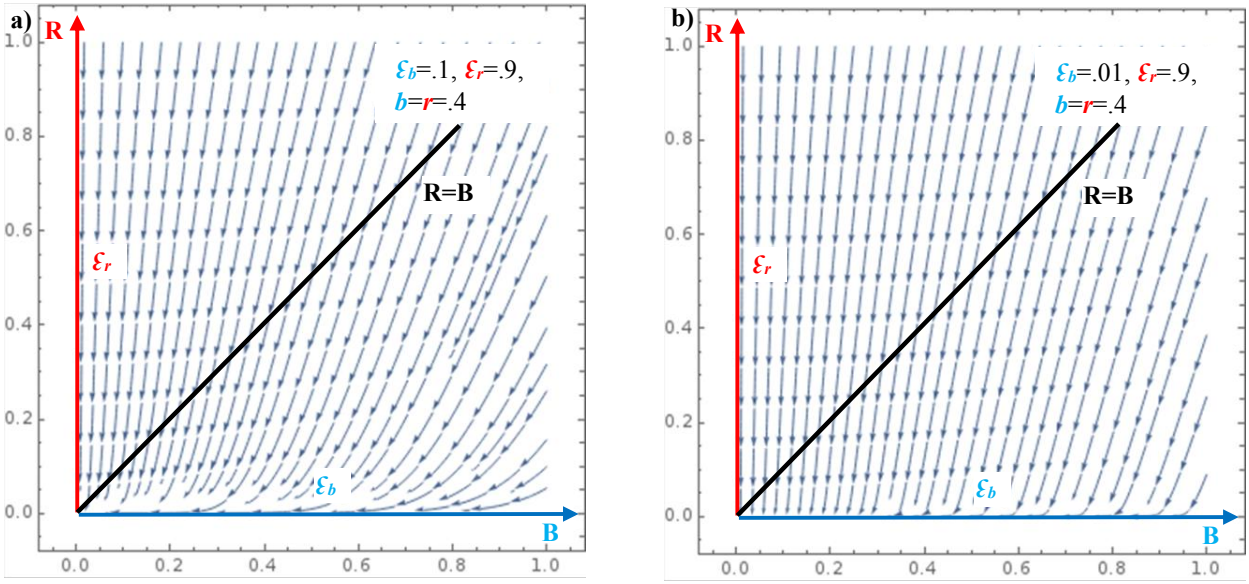


Fig. 2.6: Phase portrait of the system of equations (17): extreme eigenvalues and exposure rates

Empirically, this is easy to explain. In fact, this exactly represents the logic of ambushes or of attacks by surprise: while one side (Blue) is able to effectively locate and target the opponent's forces (Red), the Red side is not able to locate the Blue forces due to their extremely low exposure rate. In effect, one side is destroying the combat resources of the other (soldiers, equipment) at a high rate, while it almost avoids any return damage to its own combat resources.

At this point a question that was not fully explored is the impact of fire effectiveness (r and b) of the combatants on their attrition rate, independently of

their opponent's exposure rate ($\mathcal{E}_b$ and $\mathcal{E}_r$), respectively. I am providing a graphical illustration of these effects in figure 2.7a and 2.7b below.[29]

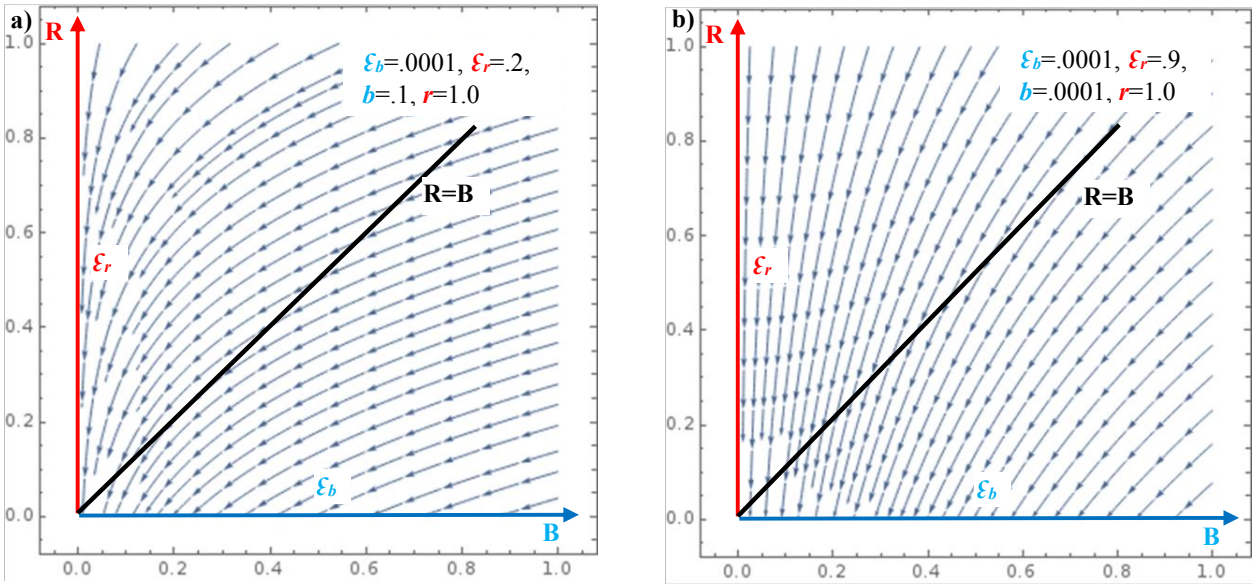


Fig. 2.7: Phase portrait of the system of equations (17): extreme fire efficiency

It is obvious that the impact of fire effectiveness (*r* and *b*) depends to a very large extent on the value of the exposure rates. From figure 2.7a we can see that Blue is facing a very high fire efficiency of Red, while responding with a fire ten times less efficient. Nevertheless, because Blue has an extremely low exposure rate this mitigates the overall dynamics in its favor. The graph reveals that initially Blue suffers some high rate of losses. However, as a very low eigenvalue takes effect, reducing exposure, the Blue can emerge victorious on the battlefield. This happens even when Blue is starting with 0.2 units of troops against 0.6 units of troops of Red.

[29] The two graphs are described by the following dynamic systems: $\dot{x}$=-0.0001x-1xy; $\dot{y}$=-0.2y-0.1xy (fig.2.7a) and $\dot{x}$=-0.0001x-1xy; $\dot{y}$=-0.9y-0.0001xy, respectively.

To further clarify the dependence of fire effectiveness on the exposure rate of the opponent and explain the lack of an inverse relationship let us examine figure 2.7b. Here, I created even more disadvantage for Blue by decreasing its fire effectiveness to an extreme, from b=0.1 to b=0.0001. The fire effectiveness of Red continues to be very high (r=1). However, I increased his exposure rate from 0.2 to 0.9. We can see then that even though Red has a very high fire effectiveness and Blue has a very low fire effectiveness, Blue reduces its combat losses closer to the fixed point, where the effect of eigenvalue is prominent. This tendency is visible but even away from the fixed point. In fact, by reducing exposure of Blue we have neutralized the impact of extremely high fire effectiveness of Red and are able to make Blue victorious despite Blue having a very low fire effectiveness.

This logic was theoretically clarified in the section that introduced exposure as illustrated in figure 2.1. It showed that in order for Blue to start taking casualties, it needs to both be in the range of Red's fire and, necessarily, be identified by Red. In the absence of this identification (equivalent to a very low Blue's exposure) the effect of Red's fire is minimal.

The intuition that we acquire by examining these phase portraits suggests that the level of troops' exposure determines *the speed of their attrition* in time. When we have two opposing sides with different number of troops, the side with smaller level of forces can realistically bring down the opponent to an equal force level. The quicker it can do that, the more successful can it be in defeating the opponent. This can be achieved either by reducing its opponent's ability to fight

through total destruction of its forces or by bringing it to the realization that fighting became too costly. Therefore, the idea that the *speed of attrition* is critically important - in the sense of empowering the weaker side to force its opponent cross the diagonal line where *B*=*R* and further drop down in the phase plane - is another insight that this analysis advances.

To better illustrate this intuition and the results of (27) I plot the two variables to show the attrition of Red (in orange) and Blue forces over time. I use the same exposure values of the two sides' troops, as used to generate the dynamics in the figures 2.4 and 2.5a.

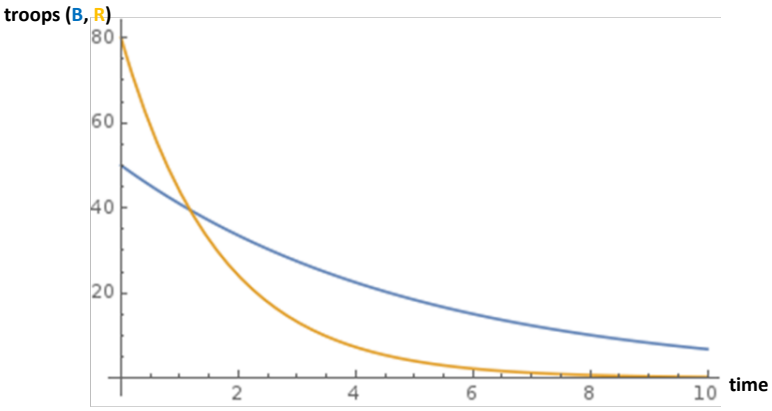


Fig. 2.8: Combat attrition of Red and Blue forces over time, (27): $\mathcal{E}_b$=.2, $\mathcal{E}_r$=.6

Based on data in figure 2.8 we can notice that the initial inputs for the number of troops (the vertical axis) were 50 units for Blue and 80 for Red. Nevertheless, due to a lower exposure rate of Blue forces ($\mathcal{E}_b$=.2) and a contrastingly higher exposure rate of Red forces ($\mathcal{E}_r$=.6), only over two time periods of interaction (the horizontal

axis) the odds of victory turned in favor of Blue forces. We clearly observe that lower exposure results in lower attrition of Blue over time, while a higher exposure of Red forces leads to almost complete degradation in their numbers already when approaching the eighth period. Shortly after the first period of interaction the Blue draws the Red forces to an equality in capabilities where the two curves cross. From the point of view of any combatant, the ability to keep the opponent forces smaller and weaker than own forces is essential, as an important observable criteria of combat success. It is often this observable fact that keeps policymakers continue supporting the war effort.

It is also of interest to examine what affects the value of *t* when Blue manages, through lower exposure, to bring Red down to its own troops' size (Red and Blue troops curves intersecting, figure 2.8). Using equation (27) we have the following:

$$x(t) = y(t) \;\rightarrow\quad c_1 e^{-\mathfrak{C}_b t} = c_2 e^{-\mathfrak{C}_r t} \tag{28}$$

we solve the inequality by taking the natural logarithm of both sides

$$\ln c_1 e^{-\mathfrak{C}_b t} = \ln c_2 e^{-\mathfrak{C}_r t} \tag{29}$$

$$\ln c_1 + \ln e^{-\mathfrak{C}_b t} = \ln c_2 + \ln e^{-\mathfrak{C}_r t} \tag{30}$$

$$-\mathfrak{C}_b t + \mathfrak{C}_r t = \ln c_2 - \ln c_1 \tag{31}$$

$$t(\mathfrak{C}_r - \mathfrak{C}_b) = \ln \frac{c_2}{c_1} \tag{32}$$

$$t = \frac{\ln \frac{c_2}{c_1}}{(\mathfrak{C}_r - \mathfrak{C}_b)} \tag{33}$$

This result suggests that the larger the difference in exposure between the two fighting forces is, the quicker can the two belligerents reach the point in time when their capabilities will be equal, cancelling thus one side's advantage in numbers. This result happens regardless of the initial asymmetry of capabilities. What a higher asymmetry in capabilities can do is to just extend the time until the match in capabilities is achieved, without being able to reverse it. This equation strongly supports the effect of the speed of attrition on combat outcomes and the driving influence of exposure on the attrition speed. It also implies that capability factor is nothing more but a reservoir of forces from which belligerents draw to support their war effort. When the exposure rates of the belligerents are similar, achieving equality in forces through attrition is very unlikely. This is because the superiority in numbers has the effect of a higher "carrying capacity", to borrow the term from mathematical biology. The logic though is that both sides suffer high levels of casualties and are depleting each other's forces at constant rates. Under equal exposure rates, the numerically superior force will then win only because it has a larger reservoir of forces, from which it supplies its combat casualties.

Let us try to falsify the effect of exposure by reversing it. When exposure is reversed and Blue forces are slightly more exposed than Red forces (but both are still highly exposed), then the graph in figure 2.9 displays a new illustration of the comparative attrition of the two sides' forces.

To further emphasize the impact of exposure on the degree and speed of attrition, let us examine another case. Consider that Blue forces are significantly

more exposed than the Red forces. This results in an even wider gap between the curves of attrition of the two sides' forces, as illustrated in figure 2.10. This is a typical picture of a war, involving a larger force that not only is able to explore well the benefits of reduced exposure. It is also capable of implementing the operational effects and strategy that increase the exposure of the numerically smaller opponent.

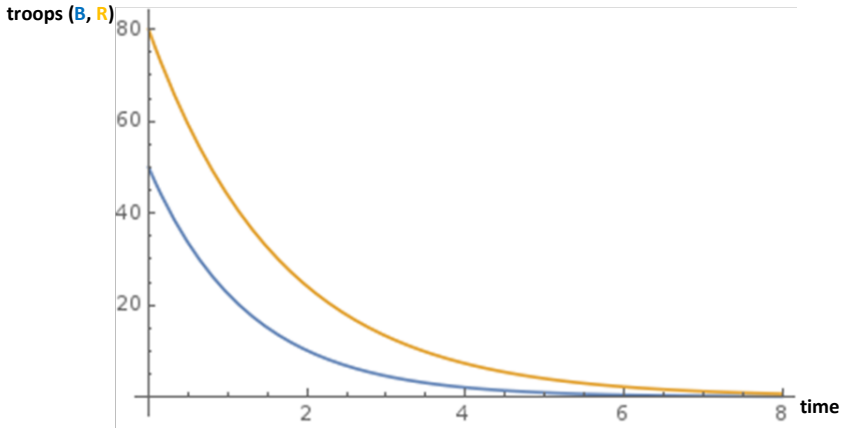


Fig. 2.9: Combat attrition of Red and Blue forces over time, (27): $\mathcal{E}_b$=.8, $\mathcal{E}_r$=.6

Empirically, such situation is possible when a technologically more advanced force can use the benefit of technology to identify the opponent and target it from distance. This is what the United States military has been doing in its military conflicts over the last few decades. The 1990 Operation Desert Shield and the 2003 Operation Iraqi Freedom in Iraq as well as the 2001 Operations Enduring Freedom are examples of such application of technologically advanced military force, allowing a combatant to reduce its exposure while increasing the exposure of the opponent.

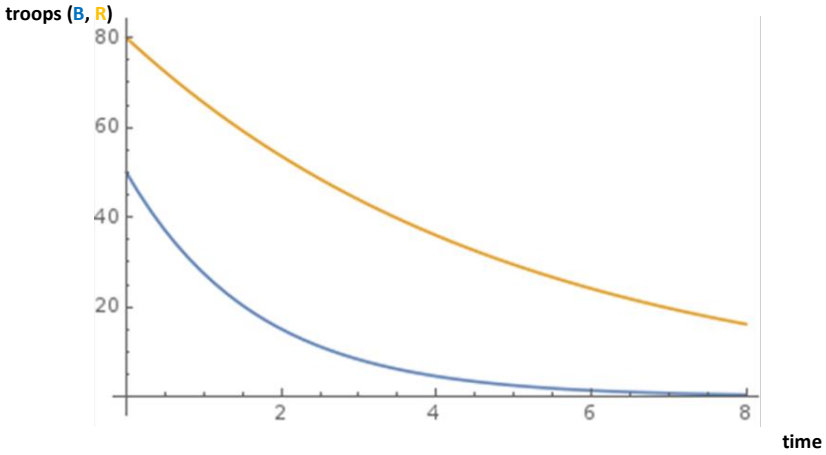


Fig. 2.10: Combat attrition of Red and Blue forces over time, (27): $\mathcal{E}_b$=.6, $\mathcal{E}_r$=.2

The 2015 Russian military intervention in the Syrian civil war has been another example illustrative of the figure 2.10 dynamics. In that conflict Russia basically copied the strategy that the United States military employed in the three above-listed military campaigns.

Let us examine a few more graphs, in order to finalize the illustration of the critical effect of exposure on the degree and speed of attrition and to emphasize the unquestionable preponderance of its effects over the effects of capabilities and fire effectiveness. Figure 2.11 presents four graphs that display the attrition dynamics of Red and Blue forces. Blue is significantly weaker in terms of capabilities but has advantage in lower exposure, from extreme to considerable levels.

In figure 2.11a Blue's exposure is $\mathcal{E}_b$=.0001 and Red's exposure is $\mathcal{E}_r$=.2. We can see that at this low degree of exposure Blue's forces do not lose almost any war-related resources, while maintaining a continuous attrition rate against Red's

troops. In figure 2.11b the exposure of Blue forces remained the same, while the exposure of Red forces was increased to $\mathcal{E}_r$=.9. As a result, we notice a considerably higher speed of attrition for Red forces. Comparing to eight periods of time in figure 2.11a, it loses almost all of its troops in some six periods of time, following a quick drop to a low number in two periods.

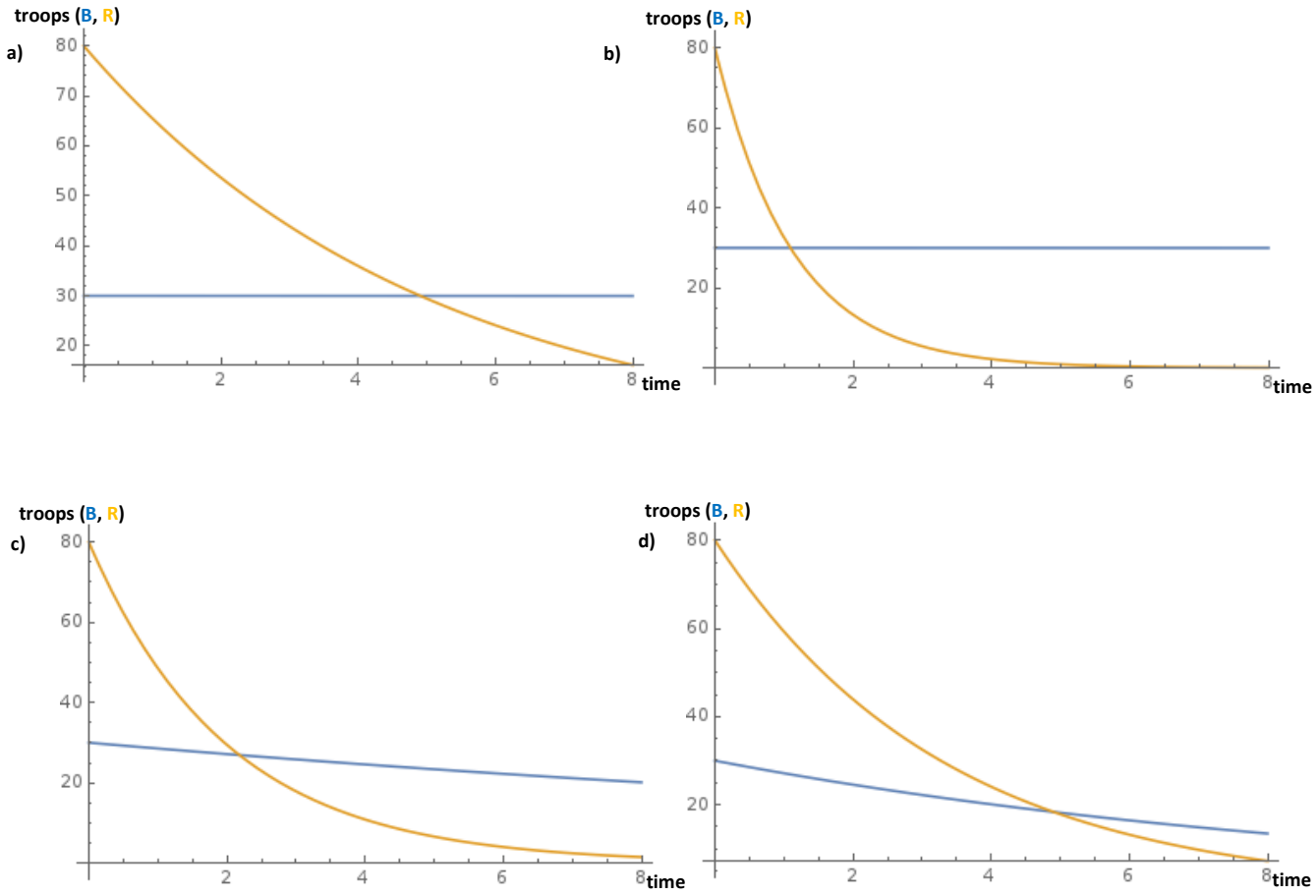


Fig. 2.11: Combat attrition of Red and Blue forces over time, (27): $\mathcal{E}_b \ll \mathcal{E}_r$

In figure 2.11c the exposure rate of Blue forces was increased to $\mathcal{E}_b$=.05, while the exposure of Red forces was decreased to $\mathcal{E}_r$=.5. In contrast to the previous two graphs the Blue forces started to take visible casualties, though at a reduced rate. However, due to much higher rate of Red attrition, Blue manages to reach an equality in capabilities already after some 2.5 periods of time.

In figure 2.11d I further increased the exposure of Blue forces to $\mathcal{E}_b$=.1, while decreasing the exposure of Red forces to $\mathcal{E}_r$=.3. The graph reveals these changes through a steeper line of attrition for Blue forces comparing to previous graphs and a lower degree of attrition of the Red forces. However, even under these conditions Blue manages to equal in forces Red after 4.5 periods of time of fighting, emphasizing the critical effect of exposure on attrition rate and speed of attrition.

In fact, keeping the figure 2.11d exposure rates for Blue and Red, the Blue troops can win even facing a 1:100 inferiority in number of forces. Under the developed model conditions, it takes some 23 periods of time for 30 units of Blue to equal in capabilities with Red that starts with 3000 units of troops. In this scenario Blue can annihilate Red troops in some 50 periods of time.

Results discussion and conclusion

The results of the current research allow us to make a few conclusions that have the potential to enrich our understanding about crisis escalation and war initiation. It suggests a novel idea about the dynamics of violent conflict and how

military capabilities map onto victory. A key factor driving the probability of victory is not capabilities or their fire effectiveness but the degree of lethal exposure of one's forces and the resulting speed of attrition.

Based on these findings we can derive that the role of capabilities is identical to that of a reservoir, from which one can supply the combat efforts. A side with higher number of forces just have a larger reservoir. If the exposure rate of combatants does not differ much, then obviously the side with a larger reservoir of war resources is likely to remain with some leftover troops. However, if the difference in exposure exists, and the side with bigger "reservoir" loses troops at higher speed, the preponderance of force advantage can very likely be undermined. Lower exposure and a resulting lower speed of attrition of forces are the two factors that affect war outcome much stronger than capabilities and their fire effectiveness.

There is an additional important intuition that this research has generated. The suggestion that exposure can affect the probability of victory and the speed of war capabilities' attrition would also imply that exposure can affect *the costs* of violent conflict. The longer two sides are fighting, the more resources they are destroying, both combat-related and those directly affecting the sides' economic development.

For instance, some sources point out that it took about 10,000 rounds to produce a casualty in World War II (Bellamy et al. 1991). Others indicate that this amount has proven to be significantly higher during the U.S. counterinsurgency

efforts in Afghanistan and Iraq.[30] Obviously, the strike force carried by a rifle bullet is not very high, based on empirical data. Research suggests that the lethality of a sword strike is higher than that of a rifle bullet. If the former has a .95 kill probability (Adamson 1977, 94), the latter has a probability kill ranging between 0.26 and 0.43, across various wars (Bellamy et al. 1991, 59-69). This micro-data indicates that modern wars tend to be much costlier in terms of expended resources to sustain the war and this is dependent on the duration of war. The level of exposure of troops can extend the duration of war (low exposure) thus indirectly impacting the costs. It can also have a direct impact on the costs, as high exposure of one side, allows its opponent to spend less munitions. Using the insights of this analysis to assess how military capabilities can transfer into victory on the battlefield can therefore also generate more accurate assessments of war costs.

This is not a challenge of existing leading research in political science. Instead, the current research is building on it by exploring new logical venues. Perceptions about military capabilities, both own and those of the competitor, play a key role in many mechanisms of crisis escalation. Existing leading research programs on the explored topic limit themselves by simply assuming the dominant role of the quantity of military capabilities. Instead, I take a modified approach and suggest looking into the micro and macrodynamics of combat interaction. This idea is triggered by the empirical fact that there is a transfer of the understanding of

[30] The Independent, "US Forced to Import Bullets from Israel as Troops Use 250,000 for Every Rebel Killed," 24 September 2005. The claims seem exaggerated, as the total number of rounds used by the troops would include training, however the number of rounds spent per insurgent killed should be still significant.

combat interaction and of what can affect military victory from the military professional to the policymaker. This encourages us to ask the question about how military capabilities or power translates into military victory.

I engage the existing rationalist approaches in political science, suggesting that uncertainty plays a significant role. However, by examining the combat microdynamics I came to the conclusion that the essential uncertainty is one determining the degree of exposure of one conflicting side's troops to the lethal effects of the other. I explore information entropy logic to show that this is the case.

An important insight resulting from this analytic exercise is that it undermines the role of capabilities in determining the probability of military victory. In fact, I show that both the superior number of forces and their higher firing effectiveness may, to a very large extent, be dependent upon the exposure of the opponent's troops, which I define and introduce. This suggests that fire effectiveness that has been one of the two dominant variables in the military operations research to explain probability of victory is in fact a proxy and is very much dependent on the degree of opponent's troops exposure. To put it differently, when troops are in direct contact, given the mass effect of modern munitions and the spatial nature of combat, fire accuracy is secondary and fully dependent on exposure.

I then explore the concept of exposure by adjusting a standard Lanchester model of combat and use it to analyze the macrodynamics of war. The results are consistent with the intuition that led to the introduction of the exposure concept. They confirm exposure's essential impact on the probability of victory. The model that I

develop provides more clarity about the combat interaction dynamics. It suggests that Lanchester model is likely accurately capturing the attrition dynamics at the point of troops contact. It is, thus, well suited to describe stationary conventional combat such as trench warfare.

However, due to the shortcomings that were widely exposed by other researchers, it fails to describe well the wider macrodynamics picture, where combat capabilities are fluid, moving constantly both on the battlefield and in the war theater. Therefore, it is not able to account for the instances when troops break contact, which introduces pauses in the attrition dynamics and even changes in the attrition intensities and direction.

This is important, particularly when analyzing modern combat. It involves a range of battlespaces including inhabited areas and that may generate additional costs for the attacker targeting them. The proposed model suggests that preponderance of force is going to be important in the instance of a total war, when many limitations in targeting various battlespaces indiscriminately will be removed. This will drastically reduce the exposure of the weaker opponent. In fact, this emphasizes the important role of uncertainty about opponent's location in combat interactions. Based on this analysis it can be claimed this is the largest uncertainty in military conflict. It is very important to be able to account for and capture this type of uncertainty. The analysis contributes to addressing the listed gaps, which have been in the attention of researchers in political science and military operations research for a while.

One way to describe what determines exposure is to refer to it as an approach in explaining the combat related uncertainty in particular, and war related uncertainty in general. Based on the information entropy framework used to clarify this related uncertainty, there are at least two significant uncertainty dimensions determining exposure: (i) the locations where the opponent is in terms of geography, fortifications, or dispersal among non-combatants, and (ii) the partition of its overall combat strength across these locations (battlespaces). While more information may be necessary and that uncertainty can be evaluated using modifications of Shannon's entropy formula, by increasing the number of questions, these two reflect the minimum information a combatant needs to evaluate the situation and derive a strategy. They provide the minimum necessary information, allowing to understand whether a belligerent's capabilities can be transferred into victory with a sufficiently high probability.

For instance, in the following chapter discussing conventional warfare, the German troops were situated on elevated positions, requiring that the British troops identify specific capabilities (heavy artillery) and design tailored for this condition strategies. In the insurgency chapter, the Chechen rebels are dispersed both through inhabited areas and uninhabited mountainous terrain. The Russian troops need to learn that distribution in order to optimize their strategies, based on the capabilities they have (no urban training) and constraints (specific sensitivities of the Russian public). Therefore, anything that increases the uncertainty about the opponent's location and partition of its forces – rough terrain, urban area making

indiscriminate targeting costly, jungles, large areas, lack of timely intelligence on opponent's numbers and weaponry, among others – has an impact on the level of that opponent's exposure.

This brings us to the understanding of how exposure can be strategically manipulated by the belligerents. They can spread disinformation, explore locations and terrain that are harder to monitor by the opponent, fortify their positions, move frequently so that the intelligence the opponent may get becomes outdated quickly, develop weapon systems that have a higher range than the weapons of the opponent, develop stealthy weapon systems, explore technical sensors, distribute their troops across terrain to increase the chances of identifying the opponent, etc. These make exposure a highly strategical concept, as it can be claimed exposure is at the heart of strategic interaction of belligerents at the tactical, operational and even strategic levels of warfare.

Therefore, this research offers additional clarity to the multiple arguments about the role of strategy or force employment in military conflicts. When claiming that strategy or other things, like terrain, are indicators of war victory, we should account for the fact that the sides use these factors to decrease their exposure. They are just tools that help materialize the effects of exposure primitive of combat dynamics. Either the strategy or the terrain can be explored to affect the degree of troops exposure and thus the speed of their attrition. This would mean that a weaker side may have higher chances to survive and even prevail, when operating in terrain that helps diminish exposure, such as mountains, jungle, and more recently, heavily-

populated urban areas. This conceptual clarification brought through the introduction of exposure, is consistent with a number of empirical analyses, which were referred earlier in the text.

This analysis does not claim to have fully solved the puzzle, as it is too early to have clarity about it. However, I would like to argue that it allows us to considerably advance in understanding better how exactly military force translates into military victory. And this is not due to numbers of troops and their effectiveness. Therefore, it offers a valuable contribution to the long-term debate of the role of quality vs quantity of military forces on military outcome. A conceptual difference is also brought by the choice not to differentiate between interstate and intrastate conflicts, as all these combat interactions are guided by the same combat dynamics.

The illustration of the micro and macrodynamics of military conflict that this analysis presented has also the potential to bring clarity to other related research projects. For instance, it offers the concrete tools to examine the role of resolve on military outcome, which is an area that continues to lack conceptual and empirical clarity. The many claims about the role of resolve in military conflict lack a solid theoretical ground and a mechanism explaining how exactly resolve is able to affect the military outcome. I would like to argue that this analysis provides the necessary conceptual ground and the mechanism to look into this question. It did it through the assessment of the role of morale on the military outcomes in conventional wars. Generally, more research is still needed in order to further advance our understanding of the role of resolve in military conflicts.

Methodologically, the analysis suggests the need to diversify the methodological approach in political science. The preponderant use of statistical methods that focus more on the observable outcomes of military conflict, without having a clear understanding of the underlying data-generating process, may pose limitation to knowledge. Without a decent grasp of the causal mechanisms that produce the observable outcomes it is hardly possible to avoid misleading conclusions through statistical analysis. A better insight into the combat dynamics should help us greatly improve the accuracy of our empirical analysis, including by suggesting that different type of empirical data needs to be collected.

In addition to its obvious goal of contributing to academic research, this analysis is guided by a significant public policy focus, aiming to examine and consider policy-relevant and feasible ideas. During the interviews that I carried out in the summer of 2014 with experts and policymakers from NATO and EU, working on military and security issues, it became clear that the interlocutors viewed the use of military tool in cost-benefit terms. Victory in a protracted and costly conflict seemed of low utility to them. These insights did not account well for the leading political science research on conflict escalation, focused on explaining failure in bargaining in terms of uncertainty about capabilities and resolve. While these studies have incontestable value for advancing fundamental research in political science, they may not reflect accurately the concerns that push policymakers to take the decision of escalating a diplomatic conflict into a military one.

The research is also driven by the idea that the ability to transfer academic knowledge into effective public policy is a way to advance the public good. Therefore, it specifically aims to address those factors that most likely have a significant impact on policymakers in crisis escalation. In fact, it would like to address the voiced priorities in the crisis research calling for the need of "offering insights into and evidence of the inner logic of crises and of the effects of moves made by states in a situation of crisis" (Frei 1982, 12). In particular, it looks at one constituent element listed in this call, and namely the drivers that make the use of force more likely in a crisis, thus escalating it (Frei 1982, 100).

I would like to claim that it succeeded in this endeavor, as the research findings provide important insights for public policy, both in terms of military policy and war-prevention. From a military operational point of view, it suggests that it makes less sense to invest into new and more powerful weaponry. This is because higher combat efficiency would be achieved by developing evasion technics and technologies of combat, such as long-range strikes, as well as intelligence gathering to quicker and better identify the opponent's location. The model indicates that preponderance of force may have impact on probability of victory, but only in close-range stationary combat and at the price of a higher level of attrition to own forces. Due to modern technological advances, accuracy of troops is less important in conventional warfare, where non-discriminatory and mass-destruction munitions such as multiple-launch rocket systems and artillery strikes are used. Fire effectiveness continues to be useful in small-scale ambushes and close-quarter

combat of infantry, though it is also highly dependent on the conflicting sides' force exposure.

From an arms control and peacekeeping perspective, it is important to design and implement international legally-binding measures that restricts non-discriminatory munitions, the targeting of inhabited areas, while also further strengthening the existing such measures. This would lead to the decreasing exposure of a defender, which will make it costlier for an offender to wage war. Generally, the implementation of exposure-related measures should make war a more expensive tool of foreign policy and would encourage states to prefer diplomatic measures in conflict resolution over violent ones.

CHAPTER III

Conventional Warfare - the First Day of Somme

The previous chapter explored and presented the critical importance of the exposure variable in combat dynamics, in contrast to the widely-held view, emphasizing firepower (capabilities) and its accuracy. In the proposed combat dynamics model exposure showed much better results at predicting both the magnitude and the speed of combat attrition and thus of the probability of victory and defeat in war. This raises an important question: namely how can the exposure variable be instrumentalized to illustrate and explain conventional battles settings better than the existing analytic tools?

Conventional warfare is a type of conflict technology in which power preponderance has been traditionally viewed as the main predictor of military victory. The challenge is to explain the considerable subset of cases when this prediction failed, or to understand the scope conditions under which victory or defeat could occur, considering also the difference in capabilities. At the beginning of the theory chapter I already listed some of the influential literature that promote the importance of power preponderance in achieving military victory, along with some other challenging the role of superior military force in determining the outcome on

the battlefield. However, it is useful in particular to frame these logical debates differently than simply addressing the "quality" of one's military or its "better strategy", as these usually emerge as *ex post*, retrospective explanations. Our knowledge on the issue would further advance if we could say what makes strategy better or what determines the quality of the troops in terms of *ex ante* preparation, effort, and hopefully, prediction of victory. I argue that the exposure-attrition causal logic responds to that demand, which emphasizes its value as an analytic tool.

It is among the few goals of this chapter to clarify the applicability of the exposure variable in armed conflict and understand how it can impact the attrition of belligerents' forces by affecting its speed. I argue that the degree of exposure to lethal fire of the opponent has a wider applicability in explaining various theories and claims about war. Because it is an underlying mechanism variable, exposure operates by verifying internal validity of a process. Thus, it could either confirm the feasibility of a war-related claim by testing its internal validity or it could indicate its failure to verify for internal validity.

Just to show the wider applicability of exposure-attrition causal relation, I would like to explore a less obvious example, which is the so-called offense-defense theory (Jervis 1978, Van Evera 1998). The offense-defense logic implies that when offense has the advantage, it was easier to defeat the opponent and take control over its resources (territory) in comparison to the scenario of defending own resources. And, when the defense has the advantage, it was easier to protect resources than to aggressively acquire them from other actors (Jervis 1978, 187).

The advantage of either offense or defense is determined by technology and geography, according to Jervis, to which Van Evera added diplomatic factors. While the man focus of this theory is not on war outcome but on war initiation, my claim is that it still can be assisted in improving its internal validity by applying the exposure-attrition logic.

The reason for this is due to its very strongly implied assumptions about war outcomes. For instance (and in addition to the definition of Jervis quoted earlier, allowing for similar implications), Van Evera (1998, 7) refers to an offensive-dominant state of the world indicating that "when *conquest is easy*, aggression is more alluring: it costs less to attempt and *succeeds more often*. Aggressors… *win their wars more decisively*". He continues by saying that when "the defense dominates… states are also deterred from preventive war by *the likelihood that their attack will fail*, *defeated* by their enemy's strong defenses" (Van Evera 1998, 9-10). I emphasized in italics the part of Van Evera's argument where the conditional assumption and expectation about war outcome is critical for the whole theoretical claim of offense-defense logic to hold. The dominance of offensive or defensive conflict technology in a historical period or in a particular state of the world makes sense *only* to the extent it increases or decreases the likelihood of military victory. Thus, the offense-defense theorists, *implicitly*, condition the validity of their theoretical claims – conflict more or less likely to be initiated - to a specific likelihood of war outcome, when victory is more or respectively less likely. Since it is not the goal of this chapter to provide an extensive analysis of the office-defense paradigm,

I provide a few more arguments in the footnote.[31] It is also suggested that we avoid ignoring another key logical element of offense-defense theory, which is the role of technology. Technology is nothing else than a multiplier of military or war-fighting capabilities, which implicitly connects the offense-defense theory to the power preponderance logic.

Another reason why I chose to explore offense-defense theory in emphasizing the ability of exposure-attrition logic to strengthen the internal validity of war-related theories is its empirical focus. For instance, Jervis claimed that it was the dominance of defense that explained the great ability of Germany to resist France and Britain's militarily superior efforts of pushing it out of France in 1914. Given the case that I examine in this chapter – the first day of the Somme Offensive – it is only natural to suggest how my proposed analytic tool fits into the already existing arguments treating the same empirical evidence. Exposure, I claim, is also able to successfully address one of the key criticisms against offense-defense theory – the accusation of excessive complexity as well as challenges to operationalize and measure it.[32]

[31] While offense-defense theorists do not make direct statements about the role of power preponderance in favoring victory in war, they do so by implication. For instance, in addition to Van Evera's claims, Glaser and Kaufmann (1998, 54) state that "offense-defense theory addresses states' decision about whether to go to war based on their judgements about whether war is likely to be successful". They also indicate that "when defense has a large advantage, even a state that is much smaller than its adversaries may still be able to afford effective defense" (Glaser and Kaufmann 1998, 51). Lynn-Jones (1995, 689) posed that "even the most powerful states will deplete their resources if they adopt offensive strategies in a world with a large defensive advantage". He also insisted that "when there is an offensive advantage, an investment in offensive capabilities produces a military force that can defeat the force deployed by a state that has invested an equal amount in defensive capabilities (1995, 666). As we can see from these attempts to explain the offense-defense balance, there is an implicit assumption that war is more likely in an offense-dominated world because a state that invests into offense believes it is more likely to win, even against usually more powerful opponents.

[32] See Lynn-Jones (1995) for a summary of the criticism.

A third reason why I chose the offense-defense theory to illustrate the wide potential of exposure-attrition pair at strengthening the internal validity of war-related theories is its extremely high popularity in public policy debates on security and among policymakers. It has been one of the most popular academic theories among practitioners, used to explain very important policy phenomena such as security dilemma, arms racing and escalation into war, even though some practitioners criticized it for failing to show sufficient empirical support (Biddle 2001).

I argue that the theoretical framework introduced in the previous chapter is able to address and reconcile most of this criticism by offering an alternative explanation for conventional combat outcomes. It is able to explain why defense or offense may be more effective strategy, under what conditions this could happen, while hinting that the offense-defense balance debate is based on a crude concept and under-developed variables. Moreover, it is able to further clarify and elaborate on the suggested critical explanatory factors of the offense-defense balance, including Jervis' proposed technology and geography, as well as Biddle (2001) choice of military tactics and strategy. In fact, the earlier-identified global variable of exposure could serve as the atomic concept upon which technology, geography and strategy explanations may build to explain combat outcomes. The heuristic explanations that previous offense-defense balance scholarly work has been advancing reflect, effectively, some observable manifestations of exposure.

Therefore, this chapter has two goals: i) to provide an illustration of exposure variable in conventional combat with the ambition to also take steps towards

"testing" the proposed theory; ii) to instrumentalize exposure in order to explain the concrete conditions, under which superiority in military capabilities can be decisive for victory or when it matters less. It will do this by examining the 1 July 1916 events in the first day of the Somme Offensive campaign, when Britain and France tried to break through the German defensive positions. The battle on that day (and the whole Somme campaign) is emblematic both for the history of warfare and the academic research of conflict. It was one of the bloodiest combat engagements not only in the British Army's history, but also in the history of warfare. Only on the first day of Somme Offensive the British had over 57,500 casualties, of which more than 19,000 were killed (Hart 2008). It is also a classic example of trench warfare of the Word War I, which in the interpretation of scholars suggested that defense technology of conflict dominated the offense one.

Research design

The main scope of this research design is to facilitate the empirical clarification and illustration of the theoretical framework built and presented in the previous, theory chapter. A secondary aim is to introduce the conditions for a testing case, examining how well the proposed theory reflects the empirical evidence.

Given the "how" type of question explored in that thesis, the selected methodological approach will be that of a case study, in line with the guidance in Yin (2003, 6) and the approaches proposed in George and Bennett (2005). Thus, I will

present a within-case analysis exploring the detailed description of the events during the first day of the battle of Somme that took place on 1 July 1916. The nuanced and detailed presentation of the case, offered in the next section, should facilitate this effort. It will also allow to explore the logic and procedures of process tracing, in order to examine the causal mechanism between the independent variables (exposure and capabilities) and the outcome of the dependent variable (attrition) (George and Bennett 2005, 205).

This specific event is very suitable for such type of investigation given it reveals strong traits of a critical case. The Somme Offensive, generally, and the first day of the campaign, in particular, represent an extreme instance of conventional combat interaction in which the empirical puzzle captured by the research question – that force preponderance can secure victory in some cases but not in others – emerges in a rich contextual environment. The within-case analysis is also suitable due to the existence, in the example selected for analysis, of a wide range of conditional variation in the independent variables of interest as well as in the selected dependent variable. It reveals the traits of a causal mechanism at play.

The question that the previous chapter examined – how military capabilities transfer into victory on the battlefield – allowed to develop a theoretical framework and mechanism explaining variation in the degree of combat attrition, which is the dependent variable. Specifically, that theoretical discussion accounted for two independent variables used in the literature - preponderance of force and fire accuracy.

Most importantly, it introduced a new independent variable – the troops' exposure to enemy lethal fire – arguing that it is the atomic variable that can uniquely explain the variation in the level of combat attrition. Furthermore, the proposed theoretical framework suggested that the exposure variable can unilaterally and categorically regulate the impact of these other two variables on troops' combat attrition, which are widely accepted in the academic and policy literature. The logic behind the theory is that in order to increase its chances for victory on the battlefield one side must strive to minimize the exposure of its war resources (personnel, military hardware, logistics), while instantly maximizing the exposure of the competitor's war resources.

In addition to the listed independent variables, this chapter is also examining some other competing explanations of combat attrition, such as morale. This has been one of the least explored in the respective literature factor affecting combat attrition. Besides some statistical correlation studies of morale's proxy – resolve[33] – we are still uncertain about the logical mechanism guiding the effect of morale on attrition. It is the aim of this analysis to exploit its theoretical framework and the selected historical events in order to offer new insights into how morale may affect combat efficiency.

The selected case provides rich illustration and evidence of failure of preponderance of force to secure victory, of conditional effects of fire accuracy, or variations in morale – all of these in the context of diverse levels of exposure. As

[33] A representative example is Maoz (1983).

such, the case offers opportunities for falsification, for examining rival theories, which is opening the venue for assessing both the validity and scope conditions of the proposed explanation and its competitors. It even provides a certain opportunity to engage into quasi-experimental exercises, given the variation in exposure (treatment) of the different British and German units, while keeping other variables, such as preponderance of force or firepower and accuracy rather stable.

The propositions that the theoretical model generated are a few. First, it indicates that at zero exposure for the Blue forces facing Red forces with even a small level of exposure, while keeping force level and fire accuracy equal and constant, their combat interaction will lead to the attrition of the Red side with no casualties for the Blue side. In fact, across the model variations the zero value of exposure for a belligerent signifies no casualties for that side. When the Blue side has a significantly lower exposure comparing to the Red side, even if the latter is significantly superior in firepower (numerically), the Blue side will still prevail, in the sense that it is likely to remain on the battlefield after the Red troops are completely destroyed.

Therefore, in accordance with the proposed model, the variable of exposure can completely cancel the advantage of preponderance of force and accuracy, as it decreases asymmetrically. The logic is quite simple. In a combat environment the level of exposure is nothing more than the proportion of forces of one side that is actively engaged in attrition through unilateral or mutual lethal action with a proportion of the forces of the other conflict side. The sets of forces of the two

conflicting sides can have an intersection in the zero set or can overlap completely – in the former case no attrition will take place and in the latter case attrition will be totally guided by the preponderance of force factor.

Exposure, thus, can be theorized formally as being dependent on two factors: i) the degree of the intersection of the two sets of forces in the physical environment; ii) the knowledge of the sides about the materialization of this intersection. This is why exposure can be unilateral, if only one side is aware of the intersection. Practically, this can happen when one side attacks the second while the latter does not know where the attack comes from; or does not have the technical ability to target well the offender, due to either distance, impenetrable armor, or her benefit of strong defense fortifications. A fighter jet attacking a convoy of Taliban fighters that does not have anti-aircraft capabilities is a good example.

In line with this logic, by using the image of the degree of intersection of the two sets of conflicting forces, the role of preponderance of force can manifest itself only in the limits of troops' intersection, given the knowledge about it of at least one of the sides. A smaller intersection will engage a smaller number of troops. This can also have unilateral effects, in the sense that the intersection can involve a small proportion of the troops of one side with even the whole force of the other side. However, in line with the proposed theory the side with the smaller number of forces can still win, depending on the relative exposure value.

The accuracy variable can be conceptualized in terms of the ability of one conflicting side to mitigate uncertainty – a lower exposure of the opponent will result

in reduced accuracy and vice versa. Thus, at zero value of exposure even the best sniper in the world would not score a kill. And, at maximum exposure values even the lousiest shooter is likely to score.

The unit of analysis explored in this case study is the combat interaction. For instance, this will include the combat engagement of the British XIII Corps with the German regiments facing them. This choice is guided by the rationale that this is the level when it is most appropriate to process-trace attrition magnitudes and the effects of our independent variables of interest upon it. The description of the case, in the previous section, was shaped so that it emphasizes these combat interactions into separate and independent causal developments. It basically constructs a chronological narrative, presenting not only the essential outline of the general case but also of the component micro-cases of separate combat interactions. Due to this, we get the advantage of both exploring the causal path in a single case, but also examine the causal power of the variables of interest by comparing how they perform across these several micro-cases.

Therefore, the analysis will build explanations for the outcome of the cases of interest, exploring multiple weaker references rather than just one exclusively strong inference. Throughout the analysis, in addition to exploring the process-tracing method, the congruence method logic will also be used. Congruence method allows the analyst to start with a ready theory and then to assess the potential of that theory in explaining or predicting the outcome in a chosen case. It explores the relation between variance in the independent variable and the variance in the

dependent variable. After examining the value of the independent variable, the researcher makes predictions about the resulting outcome of the dependent variable, in accordance with the proposed theory. Given consistency in prediction of outcomes, one can accept the plausibility that a causal relationship may exist (George and Bennett 2005, 181).

In order to deal with the possible issue of left-out variables, the process-tracing and congruence analysis will be applied to the rival explanations as well. The approach will try to disprove the alternative explanations or reconcile them, in line with case studies research tradition. The latter will be done by examining weather competing explanations address, in fact, some different segments of a more complex longitudinal development. The scope conditions of the competing theories will be scrutinized, assessing whether they can be explained or narrowed – which is an element of the theory testing endeavor (George and Bennett, 2005)

Therefore, this chapter is practically an empirical investigation of the underlying mechanism of combat attrition, which was developed in the previous chapter. In itself, it offers a novel causal path at explaining the way military capabilities transfer into victory, presenting the factor of exposure as having higher causal priority than preponderance of force, fire accuracy, or morale.

Case narrative - the context

The idea of the Somme Offensive campaign emerged following a December 1915 conference, which was summoned by the French commander-in-chief Marshal Joseph Joffre at the French military headquarters in Chantilly. It was intended to be an element of a larger effort, conducted in 1916 by the French, British, Russian and Italian armies against the German positions in Europe and including near simultaneous attacks. This, in the design of the Allies, would have placed pressure on Germany, making it difficult to use its preferred strategy of concentrating forces against each of its opponents one at a time. The offensive's location, which took place astride the River Somme in northern France, was also chosen because it was the area where the frontlines occupied by the British and French forces merged.

The Somme area wide-front offensive was initially planned to be carried preponderantly by the French forces, with the British forces in secondary role. These plans were undermined by the vast offensive that Germany launched against French positions at Verdun fortress in late February 1916. Due to heavy losses of the French side, radical changes were to emerge in the Joffre's plan, including the diminishing of the French contribution to the joint operation. This switched the role of the main attacking force to the British, including the operation's command. It also reduced the length of front on which French forces were to attack, transferring part of it to the British forces (Prior and Wilson, 2005).

There are a few important details to consider regarding the British readiness for the Somme Offensive. First, given that Britain had a small standing army before the war, it had to raise a larger force to send to the West-European theater. It ended up recruiting volunteers across the country, which reportedly formed some 92 of the 143 battalions that the British had at Somme (Keegan 1978, 219). In fact, these "Kitchener's Armies"[34] were sent to Western European front as soon as they were formed and "went to war short of every commodity save patriotic enthusiasm" (Deverell 2005, 125).

Secondly, by the end of March 1916, only several weeks after the German offensive at Verdun has begun, the French forces lost some 90,000 in dead or wounded.

In May 1916, while visiting the new commander-in-chief of the British forces General Douglas Haig at his headquarters, Joffre insisted that the Somme operation was to start as soon as possible so that it relieves the French military of the German pressure. The two military commanders agreed, preliminarily, on 1 July date for starting the offensive (Hart 2008, 37).

Because the new British volunteer units were recruited in a hurry and due to pressure to start the offensive operation in mid-1916, they were not able to receive the proper tactical training. There were not enough non-commissioned officers and officers available for the new units.

[34] This was the reference used to the volunteer troops, which were recruited at the initiative of Lord Horatio Kitchener, the newly-appointed British Secretary of War.

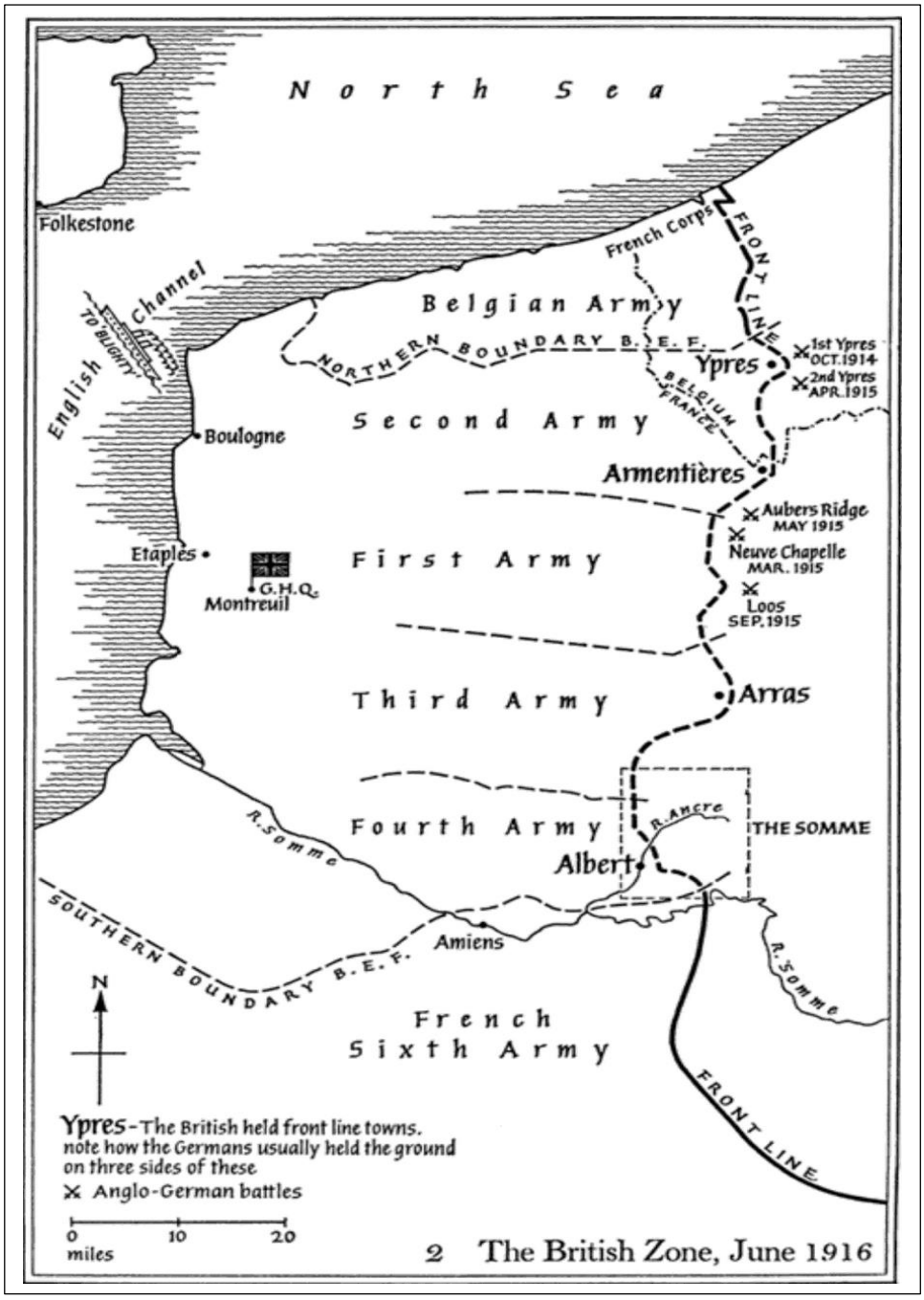


Fig. 3.1: Somme River area and the positioning of the British and French forces[35]

[35] See Middlebrook (2006).

In fact, the German evaluation of the 1 July 1916 encounter with the British troops indicated that: “The British Army… in the battle of the Somme had not yet reached a sufficiently high tactical standard. The training of the infantry was clearly behind that of the German; the superficially trained British were particularly clumsy in movements of large masses. … owing to insufficient training, they were not skillful in action” (Edmonds 1932, 490-491).

To be fair to the British forces’ leadership, various historical works indicate that Haig strongly preferred to start the offensive in the second half of the year the earliest, to also accumulate sufficient artillery capabilities (Philpott 2006, 736). As it will be suggested by the following analysis, these and a few other flaws had a critical impact on the conduct of the British offensive.

Preparation for assault

The British Fourth Army, under the command of General Henry Rawlinson, conducted a lengthy preliminary artillery bombardment of the German fortified positions, which started on 24 June. Sheffield (2015) points out that the British Fourth Army deployed 1,010 field guns and howitzers, 182 heavy guns and 245 heavy howitzers. In addition to these numbers the French units employed over 100 guns and howitzers. During the seven-day preparatory bombardment, they fired over 1,5 million shells against the German positions, of which about 1 million were

aimed to cut the barbed wire (Prior and Wilson 2005). On 1 July another 224,000 shells were fired against German trenches and fortifications (Edmonds 1932, 302).

The artillery preparation for the offensive had three objectives: to cut the barbed-wire installed in front of the German trenches, to inflict damage to the troops and material in the trenches, and to destroy the German artillery in counter-battery fire. According to Rawlinson's diary, on 26 June the weather deteriorated significantly, preventing the British forces from using their air observation to monitor the artillery impact and correct fire (Prior and Wilson 2005, 62). Due to heavy raining on 28 June the attack was further postponed for the morning of 1 July, at 7:30 am.

In line with the proposed theory, the deteriorated weather increased the level of uncertainty about the state of the enemy troops. The impossibility or difficulty to confirm the work of artillery in cutting the barbed wire was creating conditions similar to the situation when one combatant does not know the exposure level of the second combatant. However, the situation was more serious in the Somme Offensive case, as in the perception of the British commanders this was not critical. To them, the number of shells sent over the No-Man's-Land was critical instead, being the measurable indicator of success.

According to Middlebrook (2006) the wire-cutting task conducted by British artillery had certain deficiencies. Apparently, the British assigned to that task mostly their 18-pounder guns, which predominantly used shrapnel shells. Because cutting the barbed wire with this kind of shells was very difficult, requiring significant skills from the artillery crews, the task was not effectively carried out. Middlebrook also

pointed out to another deficiency of the British artillery bombardment. The British Fourth Army had an inadequate number of both heavy artillery and of ammunition for it, which were critical for the destruction of the deep dug-outs that were built by the Germans as part of their first defensive system.

Besides, due to the high pressure on the British industry to mass-produce artillery pieces and ammunition and consequently relaxed quality control by factories, it was believed that some 30 percent of the shells were duds. The consequences were severe. For instance, the X Corps lost one of their heavy howitzers on the first day of the bombardment and a second one on the next day, following a prematurely exploding shell. Historical accounts show, for instance, that one of the corps division reported 26 of its guns of different caliber had defects and malfunctioned, putting them out of action. This was a widespread challenge for all the Fourth Army corps (Prior and Wilson 2005, 62).

The artillery preparations before 1 July attack at Somme had limited effect not because the necessary tactical and operational procedures were not in place, but because these were ignored to a large extent. Edmonds (1932, 294-295) revealed that many of these procedures, which are effective measures being employed even in modern times, were known. These included cooperation between artillery and observation aircrafts of the Royal Flying Corps, as well as the cooperation between infantry and divisional artillery. The latter was maintained through forward observing officers, battery commanders' observation posts, liaison officers between artillery and infantry troops at brigade levels, and through signalers.

Troops even started to use flash spotting and sound ranging for guiding counter-battery fire.

The British failure to diminish the German defense and fire capabilities could only partially be excused by bad weather and dud shells. The British commanding staff was regularly receiving information about the results of artillery fire. During the artillery bombardment phase, the British infantry divisions sent small night patrols into the No Man's Land. They aimed to inspect the effects of shelling on the German barbed wire. Their reports indicated that considerable segments of the front had the barbed wire intact. These night raids also heard singing of the German soldiers coming from the deep dug-outs in the German trenches, contradicting the optimistic assurances of the British higher command that German infantry numbers has been drastically reduced (Middlebrook 2006). The response to the concerns, which were voiced by the front-line British infantry to their superiors, was that that attack would be a walk-over and that artillery did a great job.

If the British superior officers understood the interaction between the degree of evasion and the level of attrition, they would have paid much more attention to these reports. Because the attacking ground would to be saturated by German fire, any obstacle in the way of attacking British soldiers would have an effect of considerably increasing the British forces exposure.

In addition, many military experts and historians pointed out to the exaggerated insistence of Haig to capture a larger depth of German defense, contrary to the recommendations of Rawlinson (Sheffield 2015). The commander of

the Fourth Army was well aware of the difficulty that the task of storming the German position imposed. His initial attacking concept, sent to General Haig, included a two-step “bit and hold” approach. According to Rawlinson, the aim of the attack would not be to capture terrain but “to kill as many Germans as possible with the least loss to ourselves” (Prior and Wilson 2005, 41). One outcome of the Haig’s order was that the artillery support was spread fatally thin, and instead of concentrated fire it ended up delivering a more dispersed one. The front of planned attack was too large to be sufficiently saturated by artillery fire. In terms of the proposed theory, the British forces failed in a few areas: (i) while knowing exactly the battlespace occupied by German forces it failed to saturate it by fire - an action that would have maximized the combat attrition of these forces; (ii) given the British troops were to attack in the open, thus being forced to accept a close-to-maximum exposure value, the (i) failure led to a missed opportunity to diminish the attrition pressure on its own attacking troops by degrading the German capabilities. Or, the theory suggests that when keeping a constant and positive level of exposure, any variation in capabilities would result in a proportional variation in attrition.

Another illustrative example is the account of the Fourth Army artillery bombardment results presented by Prior and Wilson (2005). In the northern and central areas of responsibility of the Fourth Army, manned by the VIII, X, and III Corps (fig. 3.2) the patrols sent to investigate the results of artillery bombardment of the German positions brought non-encouraging news. In some areas, patrols found the barbed wire obstacles not destroyed sufficiently for them to go through. In other

areas the German soldiers used machine-gun and rifle fire to drive off the patrols. And this happened while a patrol, reportedly, managed to see a German front trench that was “totally destroyed”.

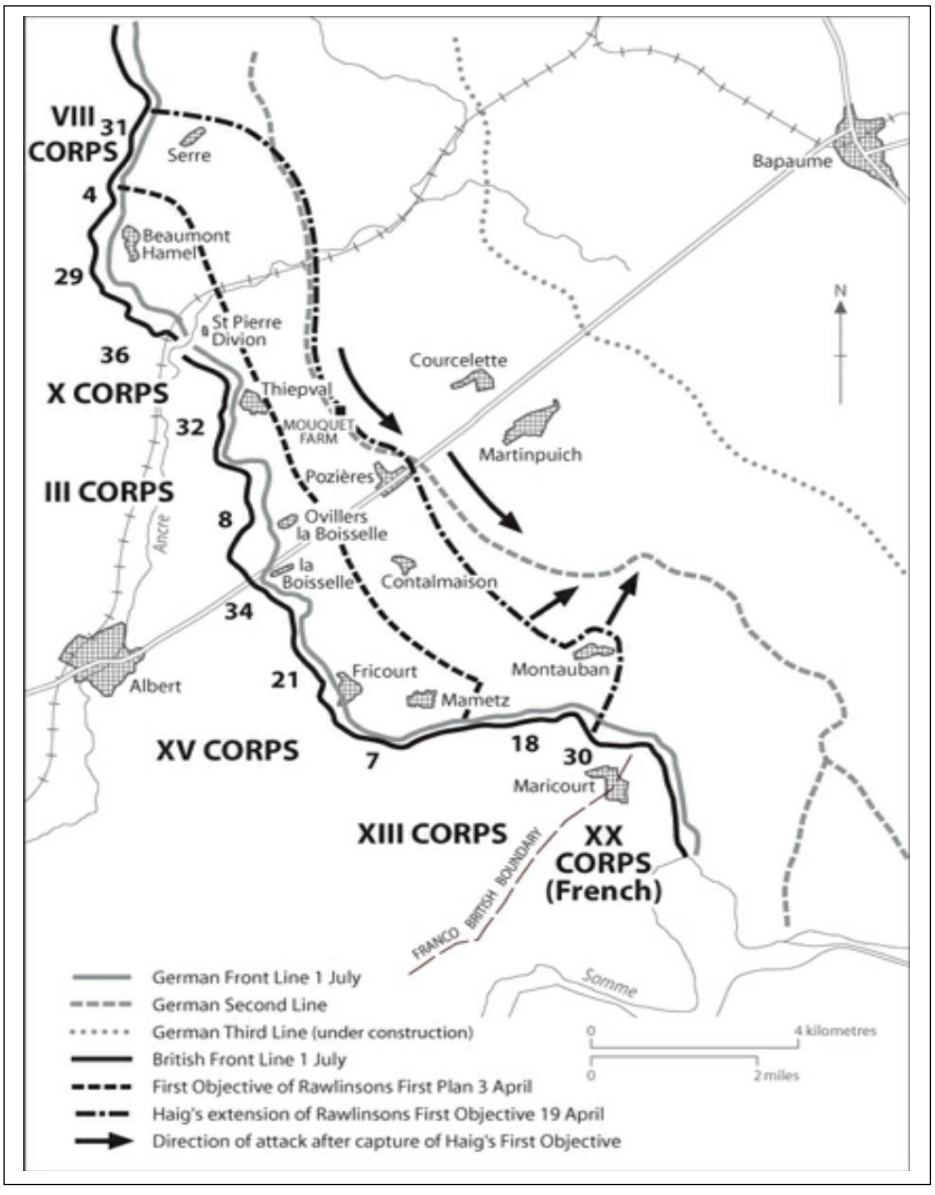


Fig. 3.2: Distribution of British corps at Somme by 1 July[36]

[36] Prior and Wilson (2005)

However, even in this case the patrol could not advance further because it was faced with “fire from a line or lines in [the rear that] very effectively sweeps the ground in front of the first line”. Prior and Wilson rightly noticed that a destroyed trench is not sufficient evidence of a successful artillery mission, if the German defenders can use it to meet the British attackers with intense machine-gun fire.

Furthermore, after examining the documents of the British Fourth Army, Prior and Wilson obtained the results of the interrogation of several German prisoners. While their numbers were insufficiently representative for the objective assessment of the issue, the information they shared does suggest that the southern part of the front might have been damaged more than the northern side. Moreover, prior to being captured, one prisoner had moved from the area opposite of the British III Corps to the area facing the XV Corps; he gave the impression that the area to the south had less robust dug-outs than his previous norther area posting.

Examining the topography of the battlefield, we can see that as the German defensive positions go south, towards Somme River, the geography in front of the XV Corps positions is less elevated (fig. 3.3) than the one in front of the III Corps. Similarly, with some few exceptions, the German positions were situated on a lower elevation across the frontline from British XIII Corps and French XX Corps.

Coincidentally, it is in the south where the British and French forces achieved the largest success in capturing their objectives and having relatively less

casualties. In fact, the French troops acquired all of their 1 July attack objectives and had modest casualties.[37]

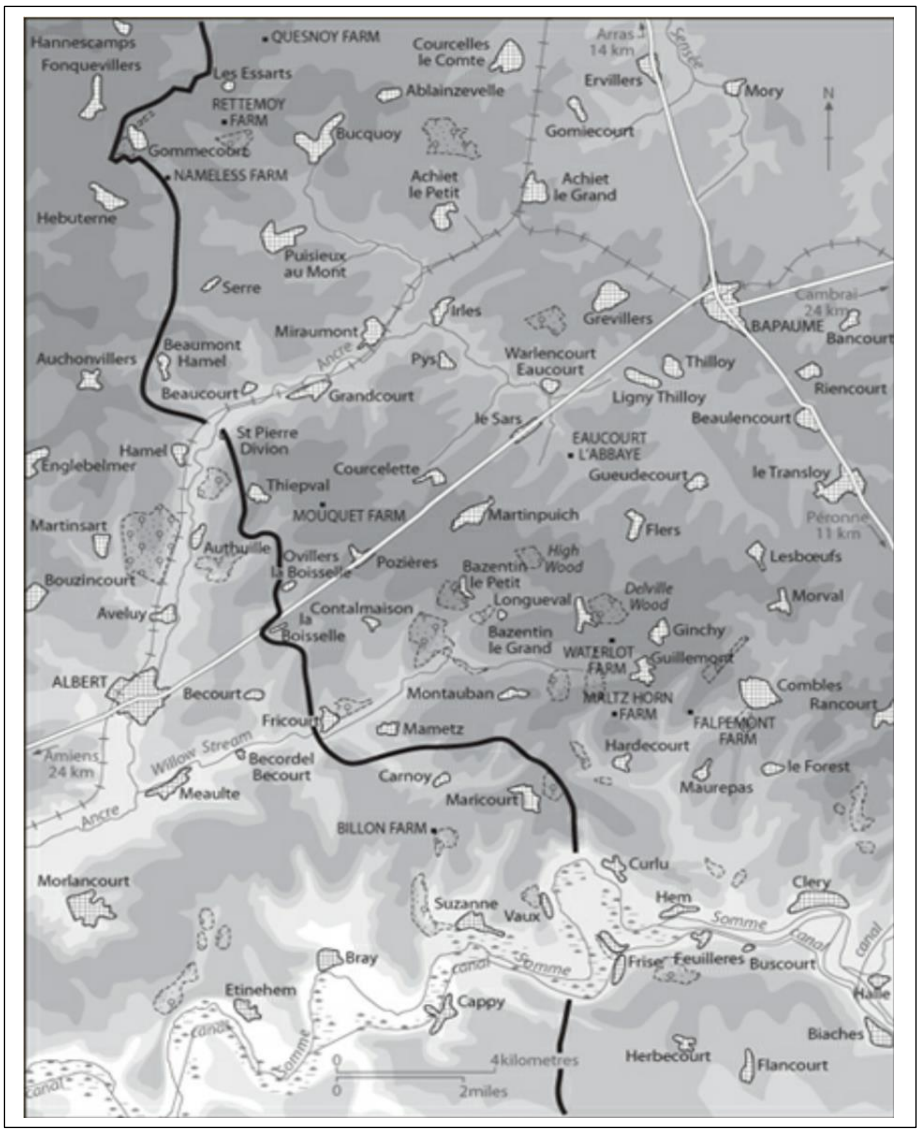


Fig. 3.3: Topographic map of the combat area at Somme River[38]

[37] The French troops also used more superior infantry tactics than the British forces, as it will be illustrated further in the text. However, the British forces of the XV and XIII corps did not copy the French tactics and explored their infantry more similar to its other corps to the northern flank of the offensive.
[38] Prior and Wilson (2005).

Moreover, if we study the map of the territorial gains that the British and French troops made after the whole period of the Somme Offensive, we do see more advance in the southern side of the front. Alone, the prisoners' responses should provoke skepticism, but along with these other details, it should at least allow us to give them consideration. The more successful British actions on a plainer battlefield, both in terms of the outcome of artillery fire and in its offensive outcomes, contrasted to the British actions in rougher terrain, corroborates the research emphasizing the role of rough terrain in diminishing the degrading of defender's forces (Fearon and Laitin 2003). It also confirms the predictions of the presented exposure-based explanation of combat dynamics.

The German defense system

The sector of the Western Front on both sides of the Somme River, chosen by the British and French forces for their attack, was defended by the 2nd German Army under the command of General Fritz von Below. If at the end of 1914 the German defenses consisted of a single trench, the situation changed by mid-year 1916, when the Somme campaign started. The German troops have built a system of defenses comprising three lines of trenches some 150 to 200 yards apart. To slow down the attackers, the Germans placed in front of their trenches two belts of barbed-wire entanglements. These were 15-20 feet wide and were fixed to dug-in iron bars to prevent their easy removal. Moreover, under their trench network the

Germans built dug-outs, which was facilitated by the area's chalky ground. Constructed some 20 to 30 feet deep and manning up to 25 men, these fortifications gave sufficient protection to their garrisons from artillery fire but the direct hit of the highest caliber. The dug-outs were also interconnected, preventing their blocking (Prior and Wilson 2005, 37-38).

In addition to these, the German side also fortified the villages from the area including Beaumont Hamel, Thiepval, Ovillers, La Boisselle, and Friscourt, incorporating them into its defense system (fig.3.4). The houses in these villages were built of solid stone, with cellars, thus able to house small garrisons with machine-guns. They were surrounded by additional trenches and barbed-wire obstacles.

All the defense positions and fortifications built by the Germans were producing, as impact, a drastically reduced exposure of the German forces. They were creating an obstacle between the Germans and the lethal fire of the British attackers - their supporting artillery, in particular. This also allowed them to build, on the way of British advance, an area that they controlled well in terms of optimizing the degree of exposure. This meant the Germans controlled the level and speed of attrition of any infantry force that stepped into that area.

Behind that first layer of the defensive system composed of lines of trenches, the German High Command instructed its troops to create a similar second layer of defense positions, over 2,000 yards behind it. The German troops understood the importance of digging trenches in reverse slope positions, in order to avoid being

directly observed from the enemy lines (Foley 2011, 475), where the terrain allowed it.

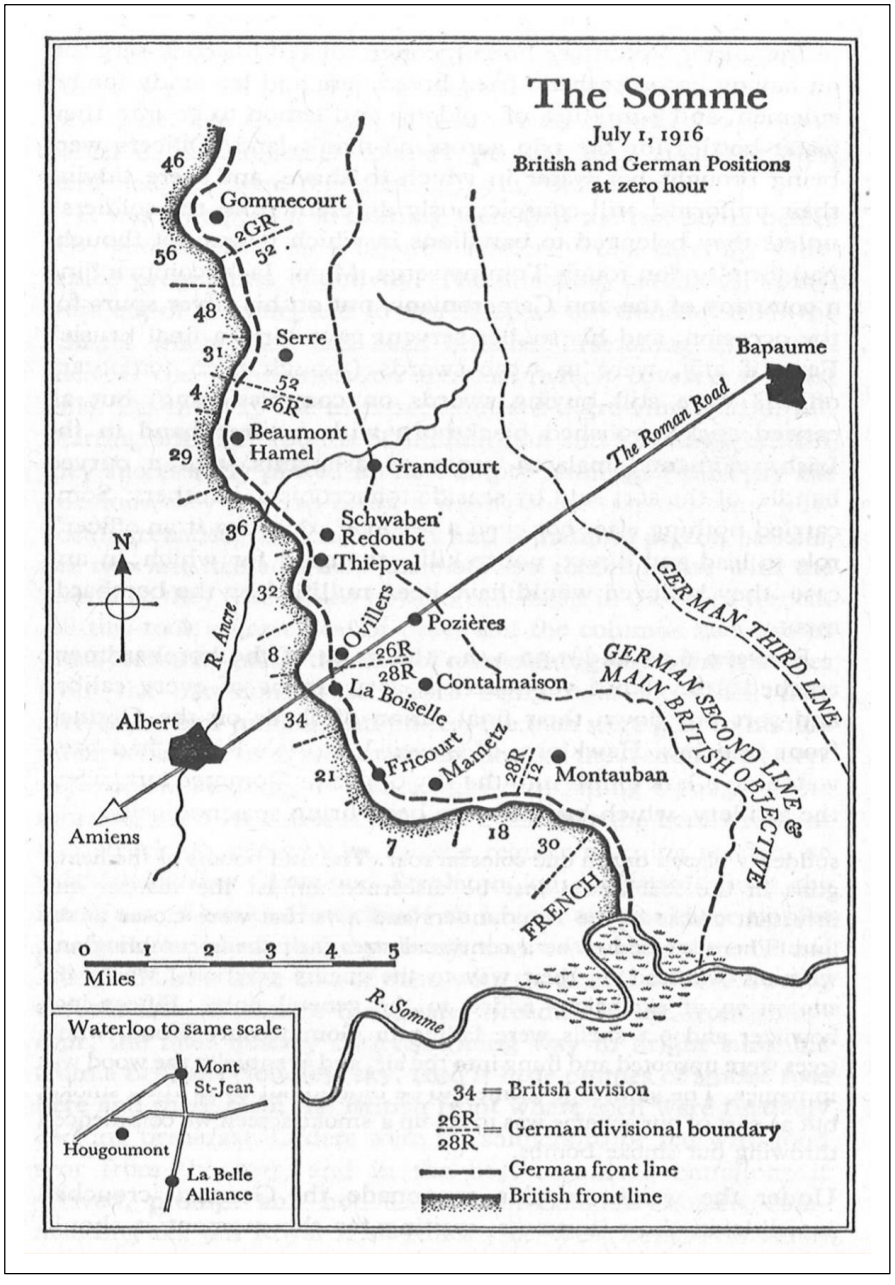


Fig. 3.4: British and German positions at Somme on 1 July[39]

[39] Keegan (1978).

However, as pointed earlier, some of the German trenches at Somme were established on the forward slopes. The Germans mitigated that vulnerability by keeping its military personnel in deep dug-outs, while deploying some observers in the trenches to warn about incoming attacks.

It was usual for the Germans to garrison two battalions of a regiment of 3,000 in or near the front system, with the remaining battalion in the intermediate strongpoints or in the second line. The data on how many German defenders of the 2nd German Army were facing British and French troops are inconsistently covered in various sources. However, Edmonds (1932, 488) suggested that the British Fourth Army had 13 divisions on its frontline, facing four and a half German divisions.

It is necessary to clarify that the British units had more personnel than the German ones. For instance, at their full strength the German infantry battalions had 750-800 men, compared to the British 1,000-strong battalion. The total fighting strength of a German infantry division comprised nine battalions. It was, thus, just slightly stronger than half of a British division (Middlebrook 2006), which consisted of twelve battalions. Therefore, a rough calculation based on Edmonds' figures would suggest the British forces had a 4.6:1 superiority in manpower.[40] It should be stressed that these calculations only considered the infantrymen and not the division

[40] Based on the data provided by Edmonds and Middlebrook, the British infantry division had around 12,000 infantrymen, while the German infantry division had some 7500 infantrymen. It has to be stressed that this does not include other types of troops in the division, such as artillery or cavalry. Moreover, at least on the German side, the size of their infantry divisions had been varying during WWI.

supporting units, including artillery.[41] The German infantry compensated inferior numbers by higher fire-power, in particular, its machine-guns.

The German side also explored various approaches to quickly employing artillery in support of infantry, when attacked. German artillery batteries practiced preparing and setting up pre-planned barrages, which could be triggered by the request of the infantry from the front-line, during the enemy attacks. It was achieved by assigning to batteries specific sectors of terrain to cover. As a result, the attackers would be confronted with a short, though intense artillery barrage (*Sperrfeuer*) in the No Man's Land, suffering significant casualties. The method was so well explored that the instructions indicated for each battery to cover a *Sperrfeuer* zone no larger than 219 yards (Foley, 476), which was not always followed.

Middlebrook (2006, 99) pointed out that during the British bombardment the German soldiers took cover in their deep dug-outs. In some sectors in the area near the Somme, the German command evacuated their whole front-line trench, hoping to get enough warning of the attack to be able to return to trenches. Technically, in sense of the developed theory, they evacuated not the trenches but a physical space where the British forces had the ability to increase the German's exposure. Thanks to the fact that Germans had advantage of higher ground, they could well see the British positions, trenches, and guns' locations. Even during the British bombardments, the German troops would deploy a sentry in the trench who would

[41] Middlebrook (2006) makes the claim that the superiority in numbers of British troops over the German troops was 7:1. Unlike Edmonds (1932) he does not provide data to clarify that claim.

observe the No Man's Land in a mirror mounted on a post; periscopes were also used by the German infantry to monitor the British trenches while in dug-outs. During the British artillery preparation, the German artillery generally held its fire, creating the impression among the British that they had been successful in destroying their opponent's artillery (Middlebrook 2006, 102).

There is excessive evidence, based on German military reports, that the German military command was aware of an attack being prepared against their Second Army at Somme. Various intelligence reports and observational data indicated of intense British and French specific activities in that area of the German frontline. However, the Chief of the German General Staff, General Erich von Falkenhayn continuously insisted, despite the strong and diverse intelligence, that the British forces attack was going to happen against the German Sixth Army position, further to the north of Somme (Edmonds 1932, 317-318).

What are the implications of this von Falkenhayn's decision, in terms of the proposed exposure-attrition theory? The decision had a smaller impact on the German positions at the northern flank of the British attack, given they were already well fortified, and thus managing to establish a reduced exposure level. However, it had critical consequences for the southern flank of the British-French joint attack. Because this segment of German defenses was not in an elevated area (which it would allow to naturally diminish German forces' exposure), and had fewer troops (meaning the Germans were able to less optimally explore the high exposure of the troops attacking them), the German defenders managed to create a lower attrition level against the British-

French attackers, resulting in more of the Allies troops reaching the German defenders. For the Germans it was much more difficult to deal with the attackers after they had left the No Man's Land with its maximum degree of combat exposure and approached the German trenches.

The 1 July attack – British and French forces

The battle of Somme on 1 July started with the following positioning of the British Fourth Army: at the most northern position was the VIII Corps, to its south immediately was the X Corps followed by III Corps, XV Corps, and XIII Corps. To the south of the Fourth Army was the French Sixth Army (fig.3.2).

The signals provided earlier by the patrols of the British Fourth Army, investigating the effects of the artillery bombardment, proved to represent quite well the 1 July patterns of attack failure and casualties of the British forces. Starting from the north, the VIII Corps' three divisions (31st, 4th, and 29th – fig.3.4) suffered horrendous casualties already in the first minutes of the attack. The VIII Corps lost about 50 percent of its overall strength, which reflected in about 14,000 casualties. For instance, the 29th Division took 5,240 casualties (Sheffield 2015), 4th Division – 5,752, and 31st Division suffered 3,600 casualties (Edmonds 1932, 447-448). Some of these divisions' battalions had casualties as high as 60-90 percent of their strength. Prior and Wilson indicated that this was due to the failure of artillery to effectively destroy any of the three key ingredients of the German defense in VIII Corps sector: barbed wire obstacles, the dug-outs and their personnel as well as

the artillery and machine guns. They offered accounts of at least 240 German guns firing on VIII Corps shortly before the 7:30 am start of attack. The 31st Division faced at least 10 machine guns targeting its advancing attackers, as another vivid illustration of the failure of the British artillery's seven-day-long preparatory bombardment. In fact, it is important to refer to Edwards (430) remarks, who indicates that the whole VIII Corps artillery, in particular the heavy one, was ordered to lift its fire 10 minutes before the assault, allowing the Germans to man their trenches.

Edwards (1932, 451-452) also provided reference to German documents, describing that some elements of defensive positions were severely degraded by the British artillery. Thus, the Germans noticed that some of the barbed-wire and many trenches were severely damaged and that British infantry was advancing very quick, to the extent that not all of the German solders managed to get out of the "great, deep-mined dug-outs, like tunnels". Due to that and the consequent British capture of a few trenches, the German forces had to counter-attack with small forces of a few platoons to push the British soldiers out of segments of the German defense positions. As a result, the German 121st Reserve Regiment facing the British 4th Division lost 560 troops, the 169th Regiment opposite of the British 31st Division had 362 casualties, and 119th Reserve Regiment that confronted elements of the British 29th and 4th divisions lost, reportedly, 292 personnel. If the figures are accurate, this would suggest an approximate loss exchange ratio of 1:10 in favor of Germans at the British northern flank on the Somme.

At the end of the day the X Corps had lost some 10,000 casualties, mostly inflicted on 36th Division – 5,104 and 32nd Division - 3,949 (Edmonds 1932). The latter was given the task to frontally attack the Thiepval village, where it was reportedly met by at least 21 machine-guns that survived the British artillery fire. Some of the 32nd Division's units were destroyed in close proximity of their own trenches (Keegan, 1978). The 36th Division, while able to go through the German barbed wire that was cut by artillery, was partly pinned down in front of enemy system of trenches by machine-gun fire; another part overran the Schwaben Redoubt (fig. 3.4), due to the destroyed trenches and their defenders in that area. However, their success was short as the German forces counter-attacked while the German artillery established a barrage of fire in No Man's Land, preventing reinforcement of and killing the withdrawing British forces. The X Corps leadership was not aware of the location of its advancing forces and tried to send the 49th Reserve Division to support the already destroyed units. Most of the division was saved from carnage because in the chaos of the situation the orders to advance did not reach it (Prior and Wilson, 2005).

The situation of the III Corps, deployed to the south of X Corps, was not much different. Its two divisions, 8th and 34th, were expected to cover a lengthy area of bare slopes until the German trenches, which were very well visible to the German forces. Moreover, the British attackers faced two villages (Ovillers and La Boisselle) across their way, situated on both sides of the Albert-Bapaume road, which were transformed by the Germans into well-defended garrisons. The regular waves of

attackers were met with machine-gun fire and artillery barrages. The German soldiers waited in their deep dug-outs for the British artillery barrage to move further towards their rear and then would come up quickly to the parapets to meet the incoming British attackers. The III Corps suffered roughly 12,000 casualties, with 8th Division suffering 6,380 and 34th Division – 5,121 casualties (Hart, 2008). This was the result of spreading the British artillery fire over a too wide variety of targets, so that the German frontline defenders were not dealt with sufficiently. Besides, due to the failure of X Corps to secure its objective, the surviving German positions turned their machine-gun fire to the south, against the flank of the 8th Division (fig.3.4). It was frequently the case that the German machine-gunners would specifically target the narrow passages through the barbed wire, as they created bottlenecks in the flow of the attackers.[42]

The XV Corps achieved most of its objectives, particularly due to much more effective fire of the British artillery, weaker German defensive positions, and shorter distance between the frontlines of the opponents. The latter allowed the British troops to quicker reach the German trenches (Hart 2008), which is equivalent to suffering a high degree of combat exposure for a shorter time. The 21st Division suffered 4,256 casualties and the 7th Division had over 3,300 casualties (Hart, 2008). The XIII Corps deployed its 18th Division and 30st Division (fig.3.4), who managed to take over the German's first line of trenches. The losses though were

[42] This is another illustration of the effect of regulating the exposure of troops, through tactical moves based on strategic interaction logic. I talk in the text about "limiting the degrees of freedom" of each other, to force an opponent to enter into a battlespace that is saturated by fire.

not small, with 3,115 and 3,011 casualties for both divisions, respectively (Hart, 2008). Similar to the situation of XV Corps, a large part of their success was due to the more effectively applied British artillery fire. In addition, XIII Corps was assisted by the French artillery from the south. Sheffield (2015) suggested that a large part of the British XIII Corps success, in particular of the 30th Division, was due to the French artillery support from its southern flank, while XV Corps benefited indirectly from the French assistance. German losses, as usually, are not fully clear. However, two regiments facing XIII Corps (109th Reserve and 6th Bavarian Reserve) had over 2,000 and 1,800 casualties respectively (Edmonds, 345). This is a stark difference from the loss ratios at the left flank of the British-French offensive on Somme River.[43]

On the top of it though, as Prior and Wilson (2005) and Edmonds (1932) pointed out correctly, the geography favored the attacks by XV and XIII Corps, as they had much better view of the German positions given a flatter topography. This allowed the artillery to be more accurate. One of the Prior and Wilson explanations for these corps' success was their exploration, however imperfect, of the creeping barrage[44], as their artillery had been accurately shelling the German trenches while the British infantry was crossing the No Man's Land. The better application of this

[43] Edmonds (1932, 341) description of the situation would suggest in the sector of XIII Corps attack a 1.5-1 loss ratio, in favor of Germans. This figure is for comparison purpose only, as given the difficulty to get exact numbers of casualties, in particular on the German side, there is some error in the provided data.

[44] A "creeping barrage" was according to Prior and Wilson (2005, 60) "one of the most important methods of infantry protection devised during the First World War. A creeping barrage consisted of a screen of shells which advanced (or crept) over the ground at a regular, predetermined pace, behind which proceeded the infantry. …It swept over all enemy defenses and was therefore likely to catch machine-gunners or riflemen whether they lurked in front trenches, communication trenches, or shell holes. Moreover, it would force dug-out dwellers to remain concealed while the attackers were traversing no man's land, or risk manning their weapons under a hail of shells. Thus, if the infantry followed close behind the barrage, they might overpower these defenders before they could emerge and bring their weapons into action".

tactic forced the German defenders to hide in the dug-outs for longer and prevented them from manning the trenches parapets quick enough, thus reducing the casualties of advancing British infantry. Other corps to the north also attempted to use this tactical procedure but lifted it too early. It resulted in German machine-gunners returning from the dug-outs to their trenches just in time to meet with fire the advancing waves of British infantry. A more likely explanation, however, is the one advanced by Edmonds (1932, 348). He suggested the XV Corps commanding officers made a good tactical judgement – by avoiding a frontal attack on Fricourt village (fig. 3.4), advancing first from the flanks, and by establishing a much better coordination between advancing infantry and the artillery, which resulted in effective fire barrage in support of the former. In terms of the proposed theory, the British had to choose between two battlespaces, one of which had a much higher exposure. By selecting the battlespace with lower exposure (the flank attack) the British managed to get more troops to the point of combat contact with the Germans, where the German advantage in exposure greatly diminished.

The disparity in losses on the first day of the Somme Offensive was significant. In addition to the already offered evidence, Keegan (1978) points out to the example when the German 180th Regiment lost 280 men out of some 3,000; while the British forces attacking it lost 5,121 out of 12,000. Overall, on the day of 1 July 1916 the British forces had 57,470 casualties of which 19,240 died, 35,493 were wounded, 2,152 were missing, and 585 were taken prisoners (Sheffield, 2015). Accurate data on the losses of German forces are not available or are contradictory.

Edmonds (483-484) indicates that the data on German casualties were only available for the duration 1-10 July, and reached 40,187, excluding the lightly wounded. Sheffield (2015, 68) advances a more narrowed-down figure, suggesting that the German forces suffered, perhaps, some 10,000-12,000 casualties, supposedly including the seven-day-period of the British artillery bombardment, in preparation for the attack. Whitehead (2013, 476) argues that some 6,226 casualties were suffered by the Germans on 1 July 1916, and an additional 1,912 became casualties due to the Allies preparatory bombardments.

The elements of the French Sixth Army operating on the southern flank of the joint British-French offensive on the Somme River seemed to have been the most successful, both in the efforts to reach its objectives and in terms of reducing its casualties. In fact, as Edmonds (1932, 342) suggests, the French forces reached all of their objectives and more. The French XX Corps, occupying positions to the south of the British XIII Corps (fig. 3.2) started its attack also at 7:30 am with two of its four divisions, employing 32 of its heavy artillery batteries. Edmonds (489) indicated that unlike the British wave-shaped attacks, the French attacking infantry went forward in small groups, exploring cover and concealment, as being witnessed by the British observation posts on the Maricourt ridge. This approach was, apparently, a part of emerging French infantry doctrine, suggesting the infantry should advance in loose formation “as skirmishers”, instead of the wave attack formation practiced widely, which was vulnerable to modern firepower (Krause, 2016).

Moreover, some of the French units were fortunate to exploit the river mist cover to reach the German front trenches unobserved. They, according to German accounts, engaged into heavy hand-to-hand combat. The French units also resisted four counter-attacks of the German forces from Hardecourt, which had apparently inflicted heavy casualties on the latter. At the end of the day the XX French Corps captured all of the German first defensive lines, without using its reserves and with limited casualties. Data on exact French casualties is not quoted in English-language literature. However, Pedroncini (1990) quoted in Philpott (2006) refers to the archives of the French Land Forces to reveal that on 2 July General Joffre noted some 4,000 French casualties for the battle of Somme, with a majority of them being lightly wounded. This data though could include the French casualties for both 1 and 2 July 1916. Given the French attack was a surprise on 1 July, it could be argued their casualties on the second day would have been bigger, lacking the surprise factor.[45]

The French forces to the south of the Somme River included the I Colonial Corps with four divisions and 65 batteries of heavy artillery, and the XXXV Corps, with three divisions and 20 batteries of heavy artillery. The German forces opposing them had reportedly only eight batteries of heavy artillery (Edmonds, 343). These two corps also did not start their assault until two hours later, at 9:30 am, thus catching the German forces by surprise. Both corps reached their objectives, though

[45] Data from a dedicated site of the French Ministry of Defense claims that French Sixth Army suffered around 1,500 casualties during the first day of the battle of Somme, https://www.cheminsdememoire.gouv.fr/fr/revue/1er-juillet-1916-offensive-sur-la-somme, accessed on 10 July 2019.

XXXV Corps suffered some higher casualties from the German forces at its southern flank. Over 4,000 German prisoners were taken by the French forces. According to German accounts, most of its artillery and its machine guns on the southern flank of the Somme Offensive were put out of action by the French and British artillery fire, while many of the dug-outs, except the deepest, were destroyed or made unserviceable. This forced many of the German forces to seek protection in shell and mine craters. These accounts corroborate the important role of terrain in affecting the combatants levels of exposure.

What affects combat attrition?

When thinking about combat attrition we mean the "wearing down" of the belligerents' combat-related resources: soldiers, military hardware, logistical and communication infrastructure, among other things. Examining the practicalities of waging war we can notice that a very basic strategic effort of belligerents is twofold: i) a first one is to limit the degree of freedom of the opponent, by forcing her combat forces into the particular battlespaces, where their attrition can be maintained at the highest possible level; and ii) the second is to instantly increase one's own degree of freedom, creating the ability to move to the particular battlespaces, where the attrition level of one's own forces can be maximally diminished. In its essence it is an optimization problem.

Because the combat interaction is strongly affected by spatial constraints, creating variation in attrition – diminishing it for one's own troops and increasing it for the opponent's troops – implies avoiding geographical spaces where the opponent can project maximum lethal fire, and moving to geographical locations from where one can project maximum lethal fire. Conventional combat is the typology of armed interactions where the degrees of freedom for both belligerent sides are minimized. If the opponent has dug defensive positions blocking a desired vector of movement, one does not have much option than to attack the opponent, which is frequently done frontally. As a rule, the opponents know the location of each other. In conventional combat the mutual exposure of the belligerents can achieve maximum values, though, as it will be clear, it can still be reduced.

On the first day of the battle of Somme the British-French forces on one side, and the German forces on the other side, knew the geographical area where the frontlines of each other was established. In fact, the German forces were deployed on a geographically elevated terrain, in particular on the northern flank occupied by the British VIII, X, and III corps. Due to this they were able to monitor extremely well the movement of the British forces. One could think about this situation as being the definitional case when exposure is absolute. However, as insightfully argued by Biddle (2004) and illustrated in his empirical studies of Taliban-US troops combat engagements in Afghanistan (Biddle, 2002), the variable of exposure of troops can be diminished even in the condition of modern conventional war.

The detailed historical narrative of the 1 July 1916 combat at the Somme River indicates that there was a considerable variation in the exposure of participating troops. This variation was noticed both across the British-French forces and across the group of German forces separately. It was the most visible, though, in the comparison between the British-French and German forces. The latter is not a surprising fact, because the German forces were dug in prepared defense positions, while the British and French forces were supposed to attack, coming in the open. Moreover, the quality of fortifications of the German side was much superior to those on the British-French side. Placed in similar conditions, when hypothetically both belligerent sides were on defense and shelling each other, the losses on the German side would have been smaller due to the better quality of their fortifications, which included their deep, 20-30 feet dug-outs in the chalky ground, able resist direct artillery fire except of the highest caliber. This was the case even when the German side was inferior in artillery, as attested by the relatively modest casualties suffered by the Germans after being exposed to seven days of British bombardment, which dropped close to two million shells on their defensive positions.

The importance of reduced exposure was illustrated in other combat engagements of the WWI. It was revealed, for instance, during the preparation for the June 1915 Souchez attack by the French. In that particular example, the German artillery counterfire managed to inflict on some of the French infantry units a level of casualties of over 10 percent of their total numbers, even before the attack was

launched (Krause 2013, 148). This was due to the fact that the French trenches and fortifications were poorly prepared. Similar examples abounded during the first day of the Somme Offensive, when complete British units were annihilated by the German counterfire, even before reaching the British frontline.

The critical role of exposure emerged, in particular, when the British troops had to storm the German defensive positions on 1 July. The chronology and details of the offensive by the British corps illustrated this. Across its frontline, and in particular in the area where British VIII, X, and III corps attacked, the British casualties were horrendous. Some of the British battalions suffered 60-90 percent losses in their manpower. Despite the spotty available data on the specific German units we can suggest loss exchange ratios of 5:1 and higher, while on some segments of the front these would reach 10:1 in favor of the German troops.

The two illustrative counterfactuals are the attacks by the French XX Corps and the joint one by the French XXXV and I Colonial Corps. The XX Corps attack was conducted, unlike the British wave-line attacks, in small infantry groups, exploring cover and by quick moves (Edmonds, 1932). The French losses were lower than the British ones by a significant magnitude, even after controlling for the effect of smaller opposing German forces. The number of German opponents though is less important, as we have numerous examples how just several machine guns were able to totally destroy a few British battalions. In contrast to the French attack tactics, which were meant to reduce the exposure of its infantry to the enemy lethal fire, the British forces were even suggested to walk at slow speed towards the

German trenches (Middlebrook 2006). Not only the wave-lines increased the exposure “surface” of the British “body” of infantry but also kept the attackers in a high-exposure battlespace, saturated with enemy’s lethal fire, for a longer duration . This had, consequently, increased the time of British forces’ lethal exposure.

The example of the attack by French XXXV and I Colonial corps initiated two hours later than the general attack is an extreme case. First, by coming later, they kept the German defenders unaware of the fact that the French attackers were in the range of German lethal fire, thus decreasing to a minimum the attackers’ exposure. Secondly, the river mist cover, exploited by some of the French attacking troops, had reduced to a great extent their exposure, as they were not noticed until reaching the German trenches. All this behavior, by exploring the natural concealment, the effect of surprise, or combat-tailored tactics of rapidly advancing in small groups, represent policy adaptations to specifically reduce the troops exposure to enemy’s lethal fire.

The power preponderance under the high exposure proved catastrophic for the British forces. The fact of them having more troops in the battlespace saturated by the German fire meant for the British side only higher losses. This argument and the explanations in the previous paragraphs go along well with the model’s predictions. Theoretically, even one teenager armed with a rifle would have been able to destroy the British Fourth Army (or the German one) if she had zero exposure and sufficient time to perform the actions. For instance, researchers suggest that most of the German casualties during the few months period of the Somme

Offensive emerged as a result of their counter-attacks (Prior and Wilson 2005, Edmonds 1932). This was the result of the policy enforced by the German General Falkenhayn, who ordered that every inch of territory lost to the opponent had to be recovered. The German counter-attacks against fortified Allies positions forced them to change roles, and accept a higher exposure of for the German troops.

Another tactical endeavor aimed to increase the exposure of the opponent on the conventional battlefield were the barbed-wire barriers. Their goal was to reduce the degrees of freedom of the attacker by forcefully increasing the time of its presence in the battlespace that was both clearly observed by the defender and was in the reach of its lethal fire. In fact, a frequent tactic of the German machine-gunners was to aim at the narrow breaches through the barbed-wire fences and gun down the lines of attackers coming through them.

Both sides explored their ability to manipulate the value of exposure, in order to undermine the effects of power preponderance, but in particular of the fire accuracy. For instance, I already mentioned the French advance in small groups towards the German trenches. In another instance, the Germans built deep dug-outs that had a high degree of resistance to British artillery direct hits, thus decreasing the German infantry exposure. The Germans also interconnected their lines of fortifications, allowing them to leave the frontline trenches during British artillery fire, effectively exiting the high-exposure battlespace.

Another tool that the military can explore to affect the exposure variable and thus the attrition of the enemy troops is the “creeping barrage”, which is a screen of

shells that moves over the ground at a regular, predetermined pace, behind which the infantry advances. Some of the quoted authors pointed out that the partial success of the XV and XIII Corps, in particular their significantly lower casualties, were the result of a correctly employed artillery barrage. The case study showed that the British corps at the left flank also attempted to use it, but (in particular in the case of VIII Corps) moved it to the second German defensive lines too early, allowing the German soldiers to man their trenches parapet and fire at the British forces, advancing in the open.

The disastrous failure of the British forces during the first day of Somme forced the military commanders to draw conclusions and learn the related lessons. In a heuristic perception of the situation, it was understood that a good cooperation between the infantry and artillery is vital for the success of an attacking operation. Griffith (1994, 66) noted that in the later stages of Somme Offensive, by 6 October 1916, an aerial observer's report was circulated throughout the Fourth Army, telling of "a most perfect wall of fire in which it was inconceivable that anything could live." Reportedly, the 50th Division followed that barrage within fifty yards, managing to take Le Sars without much opposition. The effect of the barrage, as the proposed theory explains it, is that the British artillery was able to reduce the exposure of its advancing infantry, reducing the degrees of freedom of the German forces, by separating them from the battlespace in which the British advanced and denying the Germans the ability to target it.

One less obvious way to interpret the manipulation of exposure is by thinking that the manipulating side is affecting the informational endowment of its opponent. It increases the uncertainty of the latter side about the exact location and actions of her advancing forces.[46] A more vivid illustration of this interpretation is the creation of smoke-screens by the advancing parties, either creating false barrages to make the defender believe the attack is coming from that side or aiming to camouflage its troops (Griffith 1994, 85). Of course, the smoke-screen is not a perfect protection from massive concentration of fire without specific aim by the defender in an area covered by smoke. However, it does reduce the exposure of the advancing forces by preventing the defenders from achieving more accurate targeting.

Another illustration of manipulating the value of the exposure variable and thus affecting the outcome of the attrition dependent variable is the decision by the XV Corps commanding officers to avoid frontal attack against the heavily fortified Friscourt village. The decision to use flanking maneuvers to engage the opponent where it was most vulnerable resulted in a reduction of the British forces' exposure and a lower level of their attrition.

There are also some indirect effects of exposure on the speed and degree of attrition. This is basically the other side of the story. A combatant, confronting an opponent's forces with a limited exposure and not knowing their location, could send

[46] The classical example is the one involving the attacking side's artillery "creeping barrages", when the defenders are forced to take cover to preserve themselves, thus not being able to take advantage of the attacking forces' exposure. The unsuccessful attempts by the left-flank British corps to achieve this and the more successful endeavors of the right-flank British corps to hide behind the artillery barrages were described in the case narrative.

own troops into a battlespace that is both visible and in the targeting range of the that opponent. For instance, the case study presented earlier had illustrated the way the commanding officers of the VIII Corps were not aware of the real situation on the ground. They sent reinforcements in the mistaken belief that the British 4th Division was successful in achieving its attacking objectives. Due to the lack of information about what was going on in that particular battlespace, which was severely exposed to the German fire, the British troops marched into it, and were slaughtered. This was a very frequent mistake of the British commanding officers on July 1, 1916. Assuming the first line of the German defenses were conquered, they sent reserves across the No Man's Land area, which while believed safe, was in fact completely exposed to the German machine-gun and artillery fire.

This range of examples indicates how by varying the value of exposure of troops one can affect their degree of attrition. Nevertheless, it is necessary to also examine the comparative effect of the rival explanations – the firepower, as a proxy for capabilities, and the fire accuracy - on attrition. If we control for the effect of the degree of exposure, by keeping it constant, we do see that superior capabilities may cause a higher level of attrition. Fire accuracy, as mentioned earlier, is much more dependent on the degree of exposure, provided its value is determined by the ability of the firer to diminish her uncertainty about the location of the opponent. For instance, the case study revealed vividly how the deteriorating weather did not allow the British commanders to use their air observation capabilities to correct the artillery fire and increase its accuracy.

In conventional settings the capabilities variable would manifest itself given similar or comparable-in-quality fortifications of the opponents. This intuition is reflected in the experience of the British XV and XIII Corps, which faced German defensive positions of a lower quality in contrast to the German fortifications at British forces' northern flank. In this case the superiority in artillery numbers allowed the British forces to inflict a much higher level of attrition on the German defenders. When British forces were advancing, there was very limited capacity of the German defenders to use artillery barrages against advancing British soldiers, because their own artillery was destroyed by the British preparatory strikes.

The causal path that the theoretical model proposed and the micro-comparative evidence that the explored case study revealed, had suggested that exposure completely regulates the effects of capabilities and fire accuracy on the outcome of the dependent variable – the attrition of troops.

Some critics may point out that a more numerous force is more likely to be exposed or to have a higher degree of exposure to enemy fire. This logic would ignore the definition of exposure that I employ in this study: that in order to be exposed, a combat unit needs to have its location revealed to an opponent and be in range of its strikes.

Thus, a belligerent can withdraw its forces from the opponent's range of fire, can use terrain or other cover to diminish its exposure, or can avoid its identification by the opponent while actually being in the range of opponent's fire. And, this can happen regardless of the size of its troops, in an abstract situation.

This implies that the effects of capabilities and fire accuracy are limited by some scope conditions, dependent on the value of exposure of troops to enemy's lethal fire. A zero exposure of a combat force will completely prevent the lethal effects of an opposing stronger force or of one which has higher fire accuracy rate. A very high exposure rate will create enabling conditions for the effects of power preponderance to reach its peak; similarly, this will allow to maximize the effects of fire accuracy.

A smaller force, though, if it manages to employ tactical and operational measures that keep its exposure at levels lower than that of its opponent, will be able to effectively confront a superior force. Both the empirical illustration and the theory testing attempts have confirmed the counterintuitive conclusion implied by the theoretical model, by indicating that the side with power preponderance can actually increase its attrition if the exposure of its troops is considerably higher. The case study also illustrated well how the level of exposure directly impacts on the speed of attrition and consequently on the costs of fighting. It did this by revealing the cases of British battalions being destroyed in a matter of minutes, following the impact of German machine-gun fire and artillery barrages.

The model proposed in the previous chapter (33) pointed out to the existence of a threshold, at which the asymmetric capabilities of the two sides would approach each other numerically - the speed of achieving this threshold was determined by the value of the disparity in the exposure of their troops. While the British forces' attacking waves emerged from their trenches in numbers much superior to the

German defenders, many of the attacking formations would never reach the German trenches and instead would be destroyed while crossing the No Man's Land.

What is the effect of morale on attrition?

An examination of morale effect on the level of attrition and thus on the probability of victory in war, begins with a proper definition of morale. A second step in conceptualizing it would demand the understanding of how morale manifests itself in the physical world of combat, and what is its effect on the combat dynamics. Practically, it would require imputing morale variable into the mechanism that was developed to explain combat attrition.

A separate research of morale is beyond the scope of this research. Therefore, I would refer to Peterson et al. (2008), who surveyed multi-disciplinary literature and offered a cumulative explanation of the concept. Based on multiple studies explored, they made a list of fifteen components, which included belief in group capability, enthusiasm, optimism, resilience, devotion, mutual trust, and sacrifice among several others. They suggested that the term was multidimensional and that it was a group-related concept.

Grinker and Spiegel (1945, 37) referred to morale in combat context as "the psychological forces within a combat group which impel its members to get into the fight." Manning (1994, 4) proposes a working definition of morale as "the enthusiasm and persistence with which a member of a group engages in the prescribed activities

of that group". What we are mostly interested in is how morale affects combat attrition and consequently military victory, given the physical and spatial constraints of the combat environment. What is its physical and material impact on the combat dynamics that could either increase or diminish troops' attrition? For instance, we have already seen both theoretically and through empirical illustrations that under certain conditions superior capabilities could counterintuitively lead to a higher level of attrition and increase the risk of defeat.

I would like to claim that morale enters the causal chain of interest through capabilities, as it is a quality of the human element of the combat force. If we accept this, then morale does not, independently, affect the variable of exposure. Morale, though, will instead be affected by some scope conditions, similar to those impacting the preponderance of force.

How could morale, though, affect the degree of troops attrition, while holding exposure constant and positive? Morale, based on the definition provided earlier, which mostly was developed considering the military context, would allow a combat force to confront an enemy for a longer time. In the previous chapter I discussed the Epstein (1985) research, which suggested that there exists a threshold of casualties, beyond which a fighting force would withdraw from battle trading space for time. I would argue that the physical manifestation of morale in combat is reflected in its ability to decrease this threshold. That is, a force with high level of morale is accepting to fight for a longer time and to accept more casualties before it withdraws. It was this logic that was reflected in Murray (2011, 89) when describing the German

army fighting in WWI after 1917 by stating that “the glue that now had to keep soldiers on the front was a disciplined culture of military obedience and obligation”. In fact, some more recent research supports the widely-held intuition that morale can insulate soldiers from the combat-related stress (Britt et al. 2013).

Let us start by examining the instances of high morale of a combat force and test the morale’s effects on attrition under changing values of the exposure variable. Troops with high morale that are exposed to high-intensity combat will be quicker annihilated. This will happen due to the resulting increase in their rate of attrition triggered by morale because it will extend their combat exposure. This is exactly what happened on the first day of the Somme Offensive, at the beginning of which the morale of the British soldiers and officers was extremely high. I already described in the case narrative how the German after-action reports noticed and recognizing that fact. All quoted authors similarly emphasized the high morale of British troops, repeatedly.

Following the seven day-long artillery strikes against the German fortified positions there were signals that the barbed wire was not cut, the German trenches were not destroyed, and the German soldiers continued to remain in their trenches in high numbers. Even though many of the soldiers and officers present on the frontline observed these with their own eyes, a considerable majority of the British troops continued to accept and trust the assurances of their high British command that their attack would succeed. It was also the components of morale, earlier listed, that gave energy to the waves of British soldiers to march into the field of fire of the

German machine-guns and artillery. The argument, therefore, is that the high troops' morale would very likely lead to higher attrition and increased risk of defeat in combat, under the conditions of high combat exposure of these troops.

In a second scenario, under the condition of high morale and smaller exposure of troops, the effect of morale will again maintain troops in combat for a longer duration. Like in the previous scenario, this will then be a race between the fighting sides' withdrawal thresholds values. However, reaching these threshold values depends first of all on exposure. It is when we keep exposure equal and positive for both belligerents that their speed of attrition starts to be determined by their capabilities and fire accuracy. We would be interested to consider two particular ramifications of this case: the examined actor has either superior capabilities or inferior capabilities. It is difficult to find a fighting force with superior capabilities that has lower morale, generally. Preponderance of force would trigger half of the 15 elements from the extended definition of morale, offered at the beginning of this section. Therefore, under this scenario's first ramification, it is the capabilities (through firepower) that would have the highest effect on the speed of achieving the withdrawal threshold rate.

Under the proposed model, higher morale would then explicitly manifest itself when a fighting force has inferior capabilities and which, under equal and positive exposure, will therefore have a relatively quicker speed of attrition. Having a lower threshold of withdrawal, due to higher morale, could hypothetically increase the probability that the stronger opponent would withdraw first. This is not very likely to

happen in practice, though. Success increases morale, and a larger force confronting an opponent with *similar exposure* is likely to inflict high nominal casualties, given it has more guns trained at the opponent, increasing its kill probability. The attrition of the smaller force will be quicker, and in the expectation of the victory the stronger force is unlikely to withdraw. Therefore, under this alternative scenario ramification a positive morale's effect on victory is also dependent on exposure variable.

Let us examine another two consequent scenarios of diminished morale, either under high or low exposure of troops. Although the reduced morale is expected to slow down the attrition of the troops, it will also increase the probability of defeat.[47] This will occur, basically, due to the refusal of the low morale-affected soldiers to enter combat. They will desert the ranks, for instance. Soldiers with low morale are more likely to surrender to the advancing forces of the enemy. This happened on the first day of Somme Offensive, in particular on the French frontline, when some 4,000 German soldiers were taken prisoners.[48] The low-morale soldiers can disobey the orders of their officers to fight or to go on offensive. An illustrative example is presented by the French Army mutinies in 1917, when soldiers, including

[47] It is important to consider contradictory opinions, which indicate, after studying the German troops morale during WWI that "morale alone does not provide a sufficient explanation for battlefield defeat" (Boff, 2014).

[48] The morale of the German forces was reduced, across the whole front, due to the seven-day-long bombardment by the British artillery. The German soldiers did not eat for days, reportedly, as their communication lines with the rear were destroyed and were continuously under artillery strikes. However, it was particularly at the southern flank of the joint British-French offensive that such large numbers of German soldiers were forced out of their first lines of defense or were caught by surprise by the quickly advancing French and, to a lesser extent, British forces. The listed authors pointed out to surrendering German soldiers across the frontline, though in much smaller numbers and in less frequent cases at the northern flank and in the center, in comparison to the southern flank of the British-French offensive. A significantly higher decrease in German soldiers' morale could be observed by the end of WWI, during 1917-1918, as Ferguson (2004) suggests.

veterans, conditionally refused to go to the frontline (Smith, 1995) or attack (Gilbert and Bernard, 1959). This crisis of the French military reportedly involved over 50 French divisions, close to a half of French active troops at the time (Rose 1982, 564). Rose (1982, 565) also wrote about a 1917 mutiny among British troops at Etaples in France, data about which was classified even at the time of that article writing.

These examples are important, because they represent serious cases of low morale among soldiers. Considering the proposed theoretical framework and its corollaries, I would like to claim that high morale does not ensure victory in conventional combat, being dependent on the value of combat exposure. Low morale though, could increase the risk of defeat – in 1917 the French army had at least half of its active units unwilling to fight. Considering the Millett et al. (1986, 37) proposed definition that "military effectiveness is the process by which armed forces convert resources into fighting power", low morale of troops would affect their military effectiveness, and thus increase the risk of defeat in battle. More formally, given that the Millet et al. definition sounds practically identical to the Hirshleifer (2000) definition of a contest success function, low morale would decrease the probability of success in combat, by reducing the warfighting effort, which is the domain (input) part of the function.

There is, though, an exception to this logic, suggested by the mathematical model of combat dynamics, proposed in the previous chapter (27). The exposure variable, reflected through the eigenvalues, has very high values in close-quarter

combat. In this instance, the main impact on the attrition, the dependent variable, would be carried out by the quantity of the troops engaged in fighting. In more modern combat, this would be representative of bayonet charges involving trench defenders and attackers. Technically, morale should play a considerable role in this particular type of combat engagement, because it is driven by psychological individual factors, such as hatred, fear and aggression (Kellett, 1982), and becomes an "energizing" factor (Britt et al. 2007, 35). It is under these conditions that physical efforts' spike can make a difference, given bayonets are handled like spears, as cold weapons. This burst of energy will not have a positive effect on shooting frequency, which is limited by the technical characteristics of the weapon. It will have a negative effect on shooting precision, since sharp shooting is advantaged by a cold temper. Human history reveals examples of "mad warriors scorning wounds and death – embody[ing] the spirit or reckless attack", like berserk warriors of the Nordic cultures and berserk-like warriors in other parts of the world (Speidel, 2002).

Implications for Offense-Defense Theory

The offense-defense balance theory (Van Evera, 1998) had quite an influence in security studies over the last couple of decades, including among policymakers. The logic behind it contributed to the emergence of one of the most ambitious arms negotiations in Europe – the Treaty on Conventional Arms Forces in Europe, which regulated the numbers of specific military hardware, considered of

"assault" or offensive type, and even its presence altogether in set geographical areas.[49] Yet, the theory has been criticized for failing to deliver sufficient analytic clarity (Levy, 1984), being based on poor assumption about security policy (Shimshoni, 1990), and being plagued by problems of empirical testing (Davis et al. 1998).

The attempts to defend that explanatory framework based on assumptions about war outcomes (Lynn-Jones 1995, Glaser and Kaufmann 1998) do not look very convincing. For instance, the claim that weapon technologies determine the relative costs of offensive and defensive strategies (Lynn-Jones 1995, 674-675), thus making it less expensive (more security per dollar invested) for states to seek security by adopting either offensive or alternatively defensive military postures and strategies, suffers from a number of flaws.[50] The simplest counterargument is that any defensive policy has an offensive component. For instance, General Joffre coordinated a joint offensive by Russian, Italian, British and French forces to ease the pressure of German actions on French defensive positions, which in a world of offense-defense balance would benefit from using only "offensive weapons". Similarly, Germany practiced the strategy of attacking its opponents one by one,

[49] The text of the Treaty, initiated between NATO and Warsaw Pact countries, can be found on the web-site of the Organization for Security and Cooperation in Europe, https://www.osce.org/library/14087?download=true. Russian Federation stopped its participation in the treaty in 2015, following unfruitful negotiations with US and other NATO countries on adapted provisions.

[50] For instance, one dollar spent by country A may generate a different financial impact on that country than country B does for itself by spending one dollar on defense. This would be influenced by domestic politics (more nationalistic versus liberal, the concrete preference of guns versus butter, etc.), the size of economy or existing strategic alliances (a NATO country can spend less on defense, which is cumulative for the Alliance, being a public good). Or, if the offence-defense balance explains better the word in the historical period before NATO, it would be necessary to evaluate more accurately the concept's scope conditions.

actively targeting one country while maintaining a defensive stance against the others. There are many similar modern examples at various levels, including strategic. A second point is that any weapon has a defensive (protection) and an offensive element: aircrafts can be used to bomb a defender's positions on offensive, or to protect one's positions against enemy's bombers or artillery and bomb the attacker to weaken its offensive capabilities. Or, an air defense system can be used for shooting down the aircrafts of an aggressor, exploring its anti-access and area denial functions. It can also be used by an aggressor, while on offensive, to protect his frontline and strategic installations against a defender's counter-attack efforts.

What is important about weapon technologies that are essential for the internal validity of the offence-defense balance, is that they can inflict damage (both against an attacking or defending opponent) and, especially, that they have an evasion potential. The latter is critical for reducing exposure to lethal attack of the opponent. For instance, self-propelled artillery has a higher evasion capacity in comparison to towed artillery. The capability to conduct long-range and accurate strikes is another example. In modern warfare, the existence of a multi-dimensional and complex intelligence, surveillance and reconnaissance system is of tremendous support in decreasing one side's exposure, while also increasing the exposure of the opponent.

It is important to understand that exposure is a highly strategic outcome, depending on a mutual minmax competition to reduce one's exposure while

increasing that of the opponent. An essential prediction of the presented model is that an actor with significantly inferior military capabilities has higher chances to prevail in an armed conflict, if she is able to get advantage in decreasing own exposure. The first day of the Somme Offensive revealed considerable variation along the frontline, when one side or another was favored, depending on the result in their unintended competition to reduce own's exposure to the lethal fire of the opponent.

In fact, an important intuition that the proposed model has revealed, and the case study has illustrated is that defense did not necessarily had the advantage. We have seen how the attacking French force managed to mark a sound victory against the defending German forces at the southern flank of the British-French offensive. This is while the offense-defense theory suggested defense technology was giving belligerents advantage in that particular historical case. The attackers did so by adopting an offensive operation of the type that aimed at maximally reducing the exposure of their troops. The French forces, as well as the British forces next to them, explored much more skillfully the artillery they had at their disposal in comparison to the British forces at the northern flank of the offensive. In fact, the case study indicated that the British units at the northern edge of the frontline made a number of severe mistakes in applying artillery and minimizing the exposure of its attacking forces. The mistakes can be summed up as (i) failing to reduce the exposure of own attacking troops by destroying the barbed wire, by suppressing German response fire, by diminishing the capabilities that fired on the attackers; and

by distributing the troops that moved through the No Man's Land in specific patterns diminishing their exposure to either concentrated or random German fire; and (ii) failing to increase the exposure of the German troops, by saturating the battlespace where Germans were well-known to be located.

Moreover, evidence was brought suggesting that the same weapons that were believed to favor defense in WWI – machine-guns and artillery – were successfully used by British and French forces to defeat the defending German forces not only in the first day of Somme Offensive, at the southern flank, but also in the consequent battles. What made the difference was the skill of the former in using these weapons and the strategy that they developed in applying them. Similarly, most of the quoted authors that wrote on the Somme Offensive pointed out to the numerous complaints of the German frontline infantry against their artillery, accusing it of not being effective and skillful enough in shelling the attacking forces. And the German artillery was considerably better than the British one in the discussed battle on the Somme River.

These ideas suggest an alternative way to strengthen the internal validity of the offense-defense theory and reformulate it. Choosing defense conflict technologies favor a belligerent only when this allows getting a considerable advantage in combat exposure, by decreasing it. Consequently, choosing an offensive conflict technology would favor a combatant when there are obstacles to the defender in getting advantages in exposure[51], the difference in exposure

[51] An advantage in exposure would be the case when one manages to achieve a lower level of exposure.

between attackers and defenders is not large, or, when the attacker manages himself to get an advantage in exposure. The latter case is less normal, because it is usually the defender that has advantage in exposure. However, empirically, this would be the case when the attacker explores more advanced weapons technologies, allowing him to strike from distances falling outside the attrition range of the defender. In this latter example though it is hardly possible to argue that offense dominates – it is rather the case of technological advancement of one belligerent against the other, which allows the attacker to generate asymmetrical advantage in exposure.

Conclusion

The model developed in this analysis and illustrated in the current chapter is offering strong support to the argument that force exposure manipulation is better able to affect the relative attrition rates and victory in war. The argument emphasizes an optimal and strategic approach to an active evasion of the opponent's lethal fire[52], although implies the need of a significant training of the combatant forces. The strategic manipulation of exposure requires, apart from strategic exploration of terrain, weapons and other environmental factors, the ability of the commander to quickly guide the movement of troops. The developed model has considerable

[52] We could think about that strategic "active evasion" in terms of maxmin and minmax strategies, through which each of the belligerents tries to evade the actions of her competitor aiming at maximizing her troops exposure; while at the same time each belligerent is making her own efforts at maximizing her competitors' troops exposure on the battlefield.

implications for the debate around factors determining combat outcomes, duration of war (equation 33), and cost of war (longer duration with high exposure could increase the costs of war) among other things.

The current chapter revealed and explored a less intuitive idea that even in conventional war, including some of its most intense forms like trench warfare, the variable of exposure can play a critical role, making the difference between victory and defeat. It pointed out that the classical candidates affecting defense and security policy, but also presently dominating the academic debate – capabilities, fire accuracy, and morale – can make an impact on the level of attrition of troops only through exposure. In fact, their effects can be blocked by manipulating the level of exposure to the minimum.

This chapter also offered an effective opportunity to emphasize the fact that military planners and commanders did not even conceive battlefield operations in terms of exposure. For instance, in the Somme Offensive campaign the British military commanders erroneously assumed that superiority in artillery and infantry were sufficient conditions to defeat the opposing German forces. While they obviously viewed attrition as a dependent variable in their strategizing, the capabilities and the resulting firepower were the key and only independent variables of interest. This practice continues in contemporary times, as modern military thinking is dominated by the idea that capabilities and factors augmenting them, or

serving as “force multiplier” in military jargon – technology and strategy[53] usually – should be the main areas of policy intervention and adaptation.[54]

While, at first glance, this may seem like just rebranding or redressing strategy or force employment in new clothes, it is far from the truth. Strategy, as employed, has been and continues to be about the firepower (capabilities) effect. There is another important observation. The first day of the Somme Offensive suggested that military commanders and policymakers controlling the national defense think in terms of attrition of the opponent. Diminishing own troops’ attrition has been a secondary matter. For instance, the US military preoccupation with reducing its military casualties came not from the realization of the role of exposure on attrition dynamics but due to public pressure to end the wars that generated the casualties. An implementation of the idea would require technical and operational adaptation of the military, based on a reviewed philosophy of war and combat.

The proposed explanation on how military capabilities can transform into victory provides additional opportunities for military strategists and planners, for defense industry, arms control and peacekeeping organizations, as well as policymakers. These opportunities stem from the fact that exposure, as an independent variable in the initial model, can be transformed into a dependent variable. The policymaker, then, could build the national defense and security by

[53] Strategy also is focused on “effective” application of firepower, as a rule. This is one of the major differences that is implied by the proposed theory: an effective strategy that would consider the effect of exposure should not only design an “effective” application of firepower – with the main focus to increase the opponent’s exposure – but also should consider measures that would always keep own’s exposure at minimum levels, allowed by the combat environment.

[54] A majority of strategic and operational documents of the United States defense establishment would suggest that type of public policies. As an example, see RAND report RR-2691 (Bonds et al. 2019).

designing policies that aim at preparing the national military for its combat exposure's minimization, as the need arises.

That is, the conceptualization of combat interaction in terms of exposure effect on attrition brings many public policy benefits. It creates a more discriminatory and policy-relevant explanatory theory, as George and Bennett (2005, 85) suggested the analysis should strive for. Practitioners have a higher ability to manipulate exposure, than they have in terms of manipulating capabilities. Besides, provided that directed manipulation of exposure can undermine the effects of capabilities on attrition's magnitude and the speed of attrition, it is cheaper for policymakers to engage the former variable rather than the latter.

A potential use for this theorization on war dynamics is in advanced defense technological applications commanded by Artificial Intelligence, such as swarms of drones. Computer-guided machines are better able to explore the factor of exposure, given their quick control and reaction, as well as the potential to command and coordinate multiple autonomous groups operating similar to a single combat unit. Some work on swarms has been done by RAND Corporation (Arquilla and Ronfeldt, 2000). Provided it was over a decade ago, this would suggest neither the military nor the defense industry were interested in the concept or the conceptual foundation was still underdeveloped.

Since the design and application of concepts like swarming in warfare would tremendously increase the ability to affect the magnitude and speed of attrition of the fighting units, it will also make war cheaper and thus a more attractive tool of

foreign policy. However, the exposure idea illustrated in this chapter also provides insights into how to make war more costly. This can be achieved by promoting policies of arms control or of interdictions on the use of specific types of arms and ammunitions, which increase exposure.[55] It is the case with any of scientific discoveries that they can represent the proverbial double-edged sword, having the potential for either negative or beneficial social impact. However, it is certain that without a nuanced and thorough understanding of different particularities and dimensions of warfare, we are unable to design policies that would have the chance to reduce the risk of war.

The next chapter will examine the exposure variable in the context of insurgency type of war, which will allow us to better grasp the intuition behind the idea of swarming. I will apply the logic of exposure to the idea of swarming in a separate chapter discussing the public policy implications of this research.

[55] A counterargument would be that countries may not have incentives to accept such restrictions. It would be a topic for a different research, but a few ideas can illustrate the claim. If a single country, or an alliance have advantage in manipulation of exposure, they can easily emerge victorious in conventional (non-nuclear) combat confrontations. If the technology proliferates around the world, stronger states would be interested to restrict its proliferation, which should encourage their formal engagements to restrict it through treaties.

CHAPTER IV

Insurgency Warfare – Changing the Lenses

In this chapter I am examining the theoretical and empirical implications of the exposure variable on the outcome of the insurgency wars. Both for consistency purposes, but also because it lies at the heart of the selected methodological approach, it is necessary to view insurgency type of war as a dynamical system.

By instrumentalizing the exposure variable that emerged from the theoretical dynamic model of combat attrition, I specifically attempt to explain and evaluate the belligerents' combat behavior and the incentives guiding it. The focus on combat interaction as the unit of analysis is going to facilitate the comparison with conventional warfare. A starting point for this comparison is assuming that any war aims to remove the military forces of an opponent B, which serve as a (physical) obstacle in the way of the policy goal of actor A. This removal occurs through combat interaction and mutual combat attrition. The actor who first exhausts its war resources is forced to give up opposition and yield to the demands of the other actor, who managed to preserve some of its coercive capacity after the combat engagement.

Or, in the words of Clausewitz (1976, 77):

> Consequently, if you are to force the enemy, by making war on him, to do your bidding, you must either make him literally defenseless or at least put him in a position that makes this danger probable. It follows, then, that to overcome the enemy, or disarm him, must always be the aim of warfare.

It is my understanding that many of the existing analyses depart from this theoretical framework, starting to view conventional and insurgency wars as logically divergent, rather than two variations of the same phenomenon. For instance, Arreguin-Toft (2001) asked whether the asymmetric conflict has a unique dynamic. His view is largely shared in the discipline.

Despite this different understanding of the two ways of waging war[56] there are more similarities between them than differences. With reference to Hirshleifer's (2000) statement that all combat reduces to attrition, at the point of contact, and building upon a rich military operations research literature (Ancker 1995, Perry and Moffat 1997), I am going to examine the insurgency warfare from the angle of its attrition dynamics. To justify this approach, I am going to apply the framework and intuition of the combat attrition model presented earlier and explore the impact of exposure on the incentives of both the insurgent and counterinsurgent. This endeavor, I expect, might place some pressure on a number of existing debates

[56] Some policy analysts and scholars have even argued that certain insurgencies are different from others for diverse reasons, such as losing hierarchical structure or defined goal (Hoffman 2006, Hoffman 2007), due changes in their dynamics (Berman et al. 2011, 810), including reflecting different technologies of rebellion (Kalyvas and Balcells 2010), or by having changed their means and aims (Kilcullen 2006).

surrounding insurgency type of conflicts. Nevertheless, I believe that some of my suggested contributions are complementary in analytically useful ways, and they mainly extend and clarify the domain of scope conditions affecting insurgency wars.

In my argument on conventional warfare examining the battle of Somme, I asked the reader to switch focus from capabilities, firepower and its accuracy towards a physical property affecting their performance – troops exposure to enemy lethal fire. Similarly, in the assessment of the dynamics of insurgency warfare the reader is asked to change the lenses as well. I believe we should reorient our analytic focus from examining the actors' behavioral patterns and the attempts to interpret them, towards the identification of social and physical dynamics that determine both that behavior and the actors' scope conditions.

Based on the proposed theoretical framework and the empirical examples discussed in this chapter we can derive a set of observations. First, I claim this chapter's findings suggest that the widely-popular both in the academia and policymaking dichotomy of population-centric and enemy-centric approaches to counterinsurgency is misleading and fails to completely account for the imperatives of the combat attrition dynamics. Exploring the developed model, I argue that there exists a set of imperatives that establishes a dependency of the fighting capabilities of an insurgent and even its survival on the two interrelated conditions: i) its ability to evade or diminish the degree and speed of combat attrition, inflicted by the counterinsurgent; and ii) its ability to compensate this combat attrition through extracting resources from the local population and/or foreign sponsors. In order to

survive, the insurgent would have to, at least, keep the two quantities equal. Or, if the insurgent wants to maintain any chances for success he would have to keep the attrition level smaller than the size of the compensating resource flows coming from supporters. On the side of the counterinsurgent, the secret to success is to sustain an effective *speed of attrition* of the insurgent forces, through aggressive combat engagement or interdiction methods. It has to be of a magnitude, which the insurgent cannot compensate through supplies and resources that he receives from his local and foreign supporters. It is then this specific balance between these two quantities, which can determine the violence threshold at which indiscriminate violence (brute force) becomes effective in suppressing the insurgent (Zhukov, 2014) through attrition solely. This logic significantly elevates the role and importance of combat attrition in insurgency wars, which lately, has been neglected in favor of strengthening indigenous population support of the counterinsurgent.

Another important detail that has to be emphasized is the logic that guides the supplying of the insurgent's resource flows in contrast to the aggregation of military reserves and supplies in conventional combat. If in conventional combat reserves are a part of the order of battle and they are joining the combat, as the operational needs emerge, the insurgent is by definition facing a scarcity of resources. The support it receives from the indigenous population or the foreign sponsors is being integrated into its overall capacities and operations in a much slower fashion. This specific supply dynamics, together with the insurgent's small

size, makes the insurgent particularly sensitive to combat attrition, which usually results in considerable losses of its scarce manpower and resources.

These two embedded processes reflect the main focus of this chapter's analysis and the goal to apply the exposure and attrition dynamics in explaining insurgency war.

Setting up the discussion context

The existing literature on insurgency warfare produced by both academics and policy analysts has been predominantly focusing on a binary argument – whether population-centric or enemy-centric counterinsurgency strategy is more efficient. Relatively more influential, at present, is the population-centric (or selective violence) view on insurgencies. I interpret a population-centric strategy as being one that is focusing on creating positive incentives among the population to not contribute with either of the two terms – assisting the insurgent in diminishing its exposure or compensating the insurgent for the combat losses through the provision of recruits and supplies. If these incentives are sought instead through coercive, brute-force measures, these actions are not considered part of the population-centric strategy and instead would conceptually better fit under the enemy-centric strategy. I make this analytic distinction because coercion against the population helping the insurgent can be also interpreted as the destruction by the counterinsurgent of one of the valuable resources of the insurgent.

The population-centric perspective is well known to policy practitioners through the "wining hearts and minds" metaphor. It implies that the "center of gravity" – a Clausewitz-borrowed U.S. military term indicating a critical point of a dynamic social system – is represented by the population, who is to be protected and secured (Galula 1964, Nagl 2002, Sepp 2005, Kilcullen 2009). This perspective suggests that an "iron hand" approach is rarely successful (Paul et al. 2013) and it has been consequently accepted by the U.S. Department of Defense (DoD). In its U.S. Army Field Manual on counterinsurgency (FM 3-24, 1-1) the DoD states that "political power is the central issue in insurgencies and counterinsurgencies; each side aims to get the people to accept its governance or authority as legitimate."[57] To emphasize that point, it is useful to reflect on the population-centric perspective of the top U.S. commanders in Afghanistan, since it became the U.S. military official policy and strategy in reference to that operation.

The *United States Forces – Afghanistan (USFOR-A) Initial Assessment,*[58] conducted by General Stanley McChrystal (2009) and presented to the Secretary of Defense Robert Gates called for a change in operational culture and connecting with the people. This was to be achieved by interacting "more closely with the population and focus on operations that bring stability, while shielding them from insurgent violence, corruption, and coercion" (2009, 1-2). It continued by clarifying

[57] Headquarters, U.S. Department of the Army, and Headquarters, U.S. Marine Corps. 2007. *Counterinsurgency Field Manual*. Field Manual 3-24/Marine Corps Warfighting Publication 3-33.5. Chicago, IL: University of Chicago Press.

[58] Report by the incoming NATO ISAF Commander (and Commander, United States Forces – Afghanistan (USFOR-A)) U.S. General Stanley McChrystal to U.S. Secretary of Defense Robert Gates, "Initial USFOR-A Assessment", August 30, 2009, https://www.globalsecurity.org/military/library/report/2009/090830-afghan-assessment/index.htm, accessed on August 5, 2019.

that “protecting the people means shielding them from all threats” and emphasizing that the population should be protected from insurgent coercion and intimidation. Importantly, the document also underlines that “conventional wisdom is not sacred; security may not come from the barrel of a gun. Better force protection may be counterintuitive; it might come from less armor and less distance from the population” (2009, 1-3). Now, reading this document one can grasp its strong political bias, and may instill the feeling to dismiss it in an academic research. However, this report and the consequent change in policy it triggered among the U.S. military in Afghanistan reflected well the academic work quoted earlier. Moreover, McChrystal did implement these ideas – for instance, he issued in July 2009 a tactical directive, which among other things stated that “the use of air-to-ground munitions and indirect fire against residential compounds is authorized under very limited and prescribed conditions.”[59] His successors occasionally had a similar approach, restricting the use of close air support of U.S. troops conducting missions around Afghan villages.[60] Therefore, as these examples suggest, the U.S. military has lately revealed a stronger focus on “protecting” the population, even at the expense of own force protection considerations, scarifying the lives American soldiers.

Academic experts advancing compatible ideas have applied the terms of selective and indiscriminate violence, respectively – with a majority arguing for the

[59] ISAF Tactical Directive, 6 July 2009, www.nato.int/isaf/docu/official_texts/Tactical_Directive_090706.pdf, accessed on January 14, 2014.
[60] BBC, “NATO in Deadly Afghan Airstrike in Logan Province,” 6 June 2012, http://www.bbc.co.uk/news/world-asia-18340140, accessed on January 12, 2014.

inefficiency of the latter in fighting insurgencies (Mason 1996, Valentino et al. 2004, Kalyvas 2006, Kalyvas and Kocher 2007, Condra and Shapiro 2012). However, some researchers have posed that indiscriminate violence reflecting a more enemy-centric approach[61], in contrast to the selective or population-centric one, may be a better suitable tool in suppressing insurgencies. Among them Lyall (2009) examined Russia's brutal tactics in Chechnya and showed a decrease in insurgent violence following the artillery bombardments by Russian military of villages believed to be insurgency's strongholds. Certain evidence, in this regard, was also brought by Biddle et al. (2012), who argued that it was the contribution of the 2007 surge to the so-called "Anbar Awakening" that led to a decrease of insurgent violence in Iraq. A more nuanced logic was presented by Zhukov (2014), who showed that insurgent violence can be effectively diminished by government responses with suppressive violence, if this can be elevated over a certain intensity threshold. While probably not intending so, some support to the use of suppressive strategy was lent by Berman et al. (2018, 152-183), who found that aid efforts have contributed to a decrease in insurgent violence in Afghanistan only when coupled with government suppressive strategies.

The selective violence argument tends to explore in support of its logic the rationale of information endowment effect. Most recently, this idea was strongly promoted by Berman et al. (2018), who emphasized the key role of intelligence

[61] For clarity, based on Russia's view and use of it the enemy-centric strategy is just one ignoring or allowing for high collateral damage among civilians. It makes this dichotomy analytically comparable to the indiscriminate-vs-inflammatory violence framework.

about insurgents that can be obtained from the population in order for the counterinsurgent to be successful. Preceding them with a conceptually-similar logic, Zhukov (2014) pointed to the instrumental use of violence, which he views as a substitute for intelligence.

While different threads of research preceding this analysis revealed knowledge about insurgencies responding to a "what" type of question, I am offering clarifications to them through answering the related "how" (identifying the underlying dynamics) and "when" (deriving the scope conditions) types of questions. We know that the insurgent tries to avoid full contact with the superior firepower of the opponent, by manipulating the degree of its combat exposure. For a better understanding of insurgency war, it is important to also know how exactly this is done. Among other things, an important reason in this regard is that the process of warfighting generates itself additional rational incentives to adjust the belligerents' preferences for the outcomes (Morrow 1985).

Therefore, by exploring this intuition and applying it to both belligerents, the chapter suggests that some counterinsurgents have elaborated a few adapted responses, including the attempt to increase the speed of insurgents' attrition through the application of devastating firepower. They achieve this by using *accidental* exposure of the insurgent instead of more difficult to achieve *targeted* exposure, and by attempting to reduce their own attrition speed, including through the use of indiscriminate firepower.

The rationale I am exploring in this analysis is tangential, to an extent or another, to a body of research emphasizing factors that are directly affecting the attrition dynamics. These are the arguments about the importance of the type of conflict technology applied to insurgencies (Balcells and Kalyvas 2014), the geographical factors affecting success or failures of insurgencies (Fearon and Laitin, 2003) or the idea that the "hearts or minds" approach was contextual, situational and had hardly played a role in the British military counterinsurgency design in Malaya – a logic advanced by Miller (2012). A possible take away from these studies is that the optimal strategy of the counterinsurgent will depend on specific conditions that have a directly impact on the insurgent's attrition dynamics.

I argue that different resource endowments, the combat ecology conditions, and thus the operational incentives and preferences of different, across cases, insurgents and counterinsurgents, differ. In Sri Lanka the government fought a home-grown resistance. In Afghanistan the United States is a foreign actor that is supporting the bid for power of a weaker actor, opposed by a military determined opponent. In Chechnya, the Russian Federation assessed the international image costs as less painful than what it perceived as the risk of domino effect of secessionism. There is a reason why a strategy might work in some cases but are dethroned as a successful policy in another. For instance, the proponents of population-centric explanation, claiming that indiscriminate violence does not work, would find it challenging to explain the success of Sri Lankan government's enemy-centric strategy against the Tamil Tigers (Hashim 2013, 42; Moorcraft 2012). Or,

they may find problematic the Russian eventual success in Chechnya, following a lengthy struggle to test different variations and intensities of a suppressive, enemy-centric approach.[62]

This does not signify an endorsement of policies of indiscriminate violence, which violate international law. It is just a positivist argument claiming that functionally, the indiscriminate violence would conditionally work in terms of helping a government achieve its goal. The challenge is that indiscriminate violence generates extra costs, which could take the shape of sanctions, international pressure and other diplomatic measures. Under the conditions when countries are ready to accept these costs in order to advance their preferential goals by using indiscriminate violence, we still can make a normative argument about these strategies. What I am trying to insist on, is that there is sufficient evidence undermining a functionalist argument that indiscriminate violence does not work, per se.

Therefore, in line with the intuition of the presented model I would like to argue that debates about most suitable strategy for counterinsurgency, in their current shape, are misplaced. Instead of contrasting the strategies based on superficial analysis of actor's incentives, we should focus on the attrition dynamics of the insurgent and counterinsurgent. In doing this, we shall examine, controlling

[62] There are contradictory studies conducted on the Chechen case (some claimed suppression worked, while others opposed that) and contradictory interpretations. I will explore the exposure logic to attempt to reconcile the existing contradictions, offering an alternative interpretation of the Russian counter-insurgency strategy.

for case-specific scope conditions, what actions could produce the highest variation in the belligerents' forces attrition.

As an example, let us take McChrystal's claim that the new center of gravity of U.S. military in their counterinsurgency operation in Afghanistan should be the population, even at the expense of an increased attrition of American troops. Based on the text of the McChrystal *Initial Assessment* document, we could derive a following rationale behind its claims. The document elevates as one of its new strategy's four fundamental pillars the goal to "prioritize responsive and accountable governance – that the Afghan people find acceptable – to be on par with, and integral to, delivering security" (2009, 2-2). It continues that "the population can also be a source of strength and intelligence and provide resistance to the insurgency. Alternatively, they can often change sides and provide tacit or real support to the insurgents" (2009, 2-4). The report makes a correct move by saying that the reasons of why people change sides in that conflict, must be understood. But this is largely a rhetorical device, as the document does not follow through in answering the question. Instead, it states that "ISAF and the international community must provide substantial assistance to Afghanistan until the Afghan people make the decision to support their government and are capable of providing for their own security" (2009, 2-4). Based on the exposure model this approach is suboptimal, as it does not maximize the attrition of insurgent's resources, while minimizing own's troops attrition. In fact, it encourages and triggers an inverse effect.

An attempt of analyzing the insurgency war along that logic was made by Paul et al. (2016). They suggested that the existing dichotomous population-centric versus enemy-centric debate should instead progress into a more nuanced and sophisticated 2x2 logical matrix, exploring two dimensions (fig.4.1). One of these is the space of actions, which includes both the application of physical force and political actions. Another is the space of targets, which contains both active insurgents and the insurgent support.

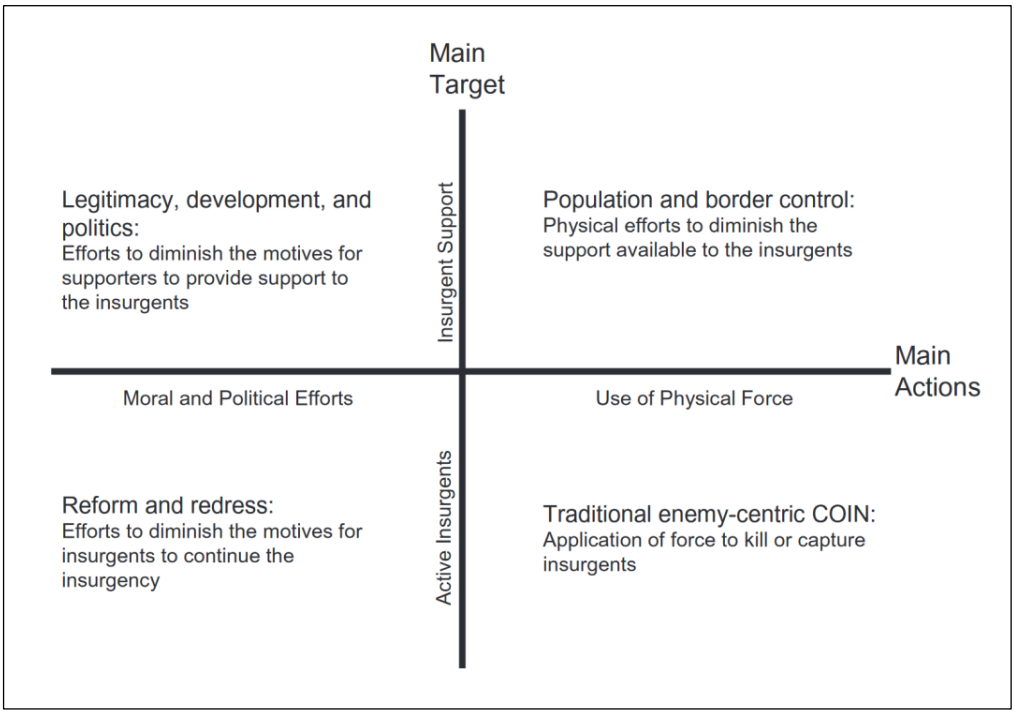


Fig. 4.1: Two-dimensional counter-insurgency model separating targets and actions[63]

[63] Paul et al. (2016, 1029).

As a result, they obtain a two-dimensional space with four quadrants: brute force is used against active insurgents (enemy-centric approach), brute force is used against insurgent support, insurgent support is addressed through moral and political efforts (a “hearts and minds” approach), and subsets of the active insurgents are addressed through moral and political efforts. Depending on the insurgency combat ecology, including social, political, and economic contexts, a state can apply policies of varying types and intensity across all four quadrants. Or, it can select only those, which are believed to be the salient drivers of insurgency – the centers of gravity that guide a tailored counter-insurgency policy.

The relative richness of this model, comparing to the dichotomous population-centric/enemy-centric, equips the analyst with enhanced explanatory power. What makes it different from other attempts to offer a counterinsurgency solution is that their model accounts for the intuition of the dynamics impacting on the degree and speed of attrition of insurgent’s capabilities (in smaller letters on the graph) although they are not explicit about the underlying logic of their proposed strategy.

Paul and his co-authors (2016) are trying to fit their model into the observed empirical regularities. In contrast, I am exploring the war dynamics model based on the exposure-attrition theory I developed, to provide a theoretical framework that has the potential to causally explain these empirical regularities. Understanding the underlying causal mechanism driving a social process is of critical importance for

improving our knowledge of that process and of its effects. It is also hard to judge the internal validity of a model, unless we understand its causal mechanism.

The analytical value of a study would emerge by clearly separating between different policies and offer insights for distinguishing among these policies' effects. Or, it would be more nuanced about the scope conditions of each approach, if a mixed strategy combining a few policies is proposed. This is what Paul et al. (2016) are trying to achieve, by suggesting the population-centric and enemy-centric approaches are elements of the same model, reflecting points on a conceptual continuum.

Another aim of this chapter is to delineate and clarify the scope conditions of various counterinsurgent responses. Alternative perspectives that explored the connection between these two supposedly competing logics have posed that information endowment plays an essential role in guiding the insurgency's dynamics and the counter-insurgent actions (Zhukov 2014, Berman et al. 2018). The chapter aims to conceptually enlarge and enrich the context in which the information argument emerges, identifying as a result a set of other relevant factors and alternatives in insurgent warfare.[64] While these and other authors imply that the combat exposure of the insurgent can be increased preponderantly through exploiting the population's benevolently-provided intelligence, the model derived by this research suggests the existence of a much wider set of possibilities for exposure

[64] The chapter is not going to explore the reverse relationship – how the insurgent affects the attrition dynamics in counterinsurgents, as this is well reflective of fig. 4.1 (c) and would be conceptually similar to a conventional combat interaction exploring a high degree of maneuver movements.

manipulation, but also of constraints for the population-centric strategy. Given exposure is a wider analytic framework that can encompass a few mechanisms affecting the dependent variable of interest – the troops' attrition – the proposed model also illuminates how to instrumentally vary attrition. This can be, for instance, achieved by extending the physical space of firepower application and switching from a targeted exposure to what I call accidental exposure. In the next section I am presenting a clarifying discussion of the microdynamics of the insurgency conflict technology.

The microtechnology of insurgency

One of the conclusions of this chapter sums up the idea that the difference between the conventional and insurgent type of warfare is one of degree, not kind. In insurgent wars we have the weaker of the belligerents deliberately choosing as its main strategy the relative reduction of its fighters' exposure. Therefore, using the combat attrition model I define an insurgency-driven conflict as one in which the weaker side tries to compensate the disadvantage in capabilities and firepower through asymmetric advantage in exposure. To see this let us examine the microtechnology of insurgency war: by observing the model's dynamics, in particular by looking at the attrition equation using x to refer to the counterinsurgent and y to - the insurgent.

$$x(t) = c_x e^{-\varepsilon_x t};$$

$$y(t) = c_y e^{-\mathcal{E}_y t} \tag{27}$$

Technically, these are decay functions, which indicate that a certain quantity of interest degrades over time. Thus, the initial level of insurgent's fighters, c_y will decrease through attrition, dependent on their combat exposure $\mathcal{E}_y$. By managing to keep the value of $\mathcal{E}_y \ll \mathcal{E}_x$ the insurgent can extend the time of military resistance against the counterinsurgent. Let us illustrate this logic though some visual aid (fig. 4.1).

It represents a schematic illustration of conventional warfare of a positional type (a) and of a maneuver type (b), along with a logical illustration of insurgency warfare dynamics (c).

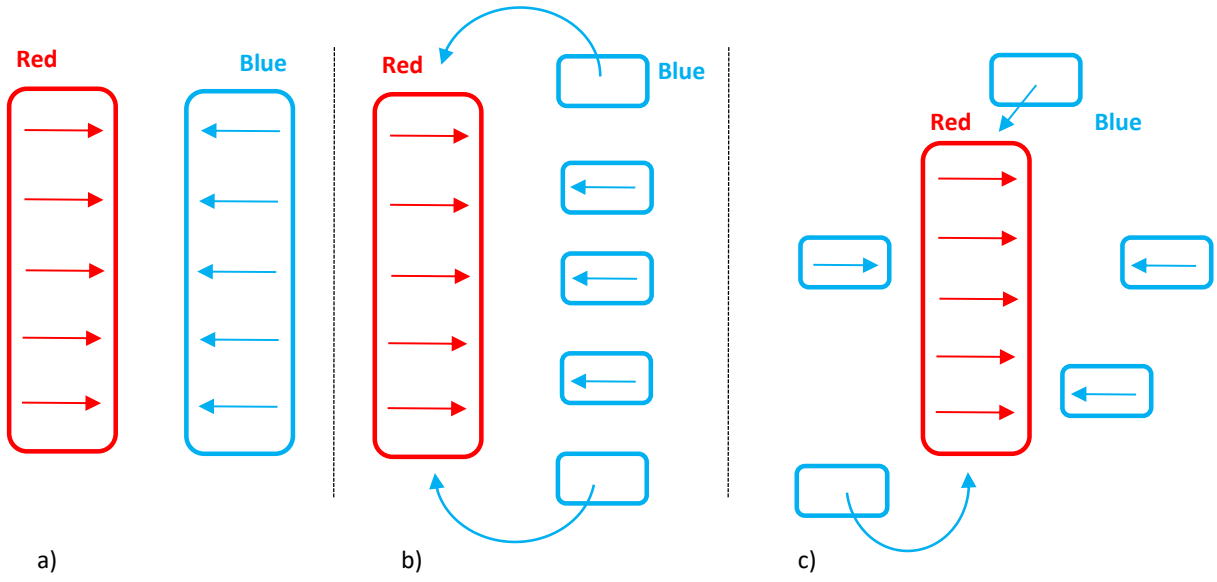


Fig. 4.2: Conventional war dynamics of positional type (a), maneuver type (b), and insurgent war dynamics (c)

Visualized in this way, we can actually see that the only difference among them is in the degree of dispersal of the Blue units, in addition to the difference in its lines (angles) of attack, contrasted to the Red troops' lines of attack.[65]

In fact, the schematics of insurgency warfare (fig.4.2, c) could very well reflect a similar situation in a conventional war when smaller units of Blue would manage to flank the Red troops. Or, we could well encounter a (c) scenario in conventional war when the advancing Red troops are ambushed by the Blue troops. An important consideration to account for is that the field of fire of the weapons is unidirectional. The significance of this is that it takes time to re-direct the arms for the attacked side, which is exploited by ambushes, maneuvering troops, or insurgents, as they start with their weapons already trained on their targets.

This is an important microdynamics trait of the insurgency conflict technology, which existing research frequently ignores since it tends to focus on the macrolevel. Therefore, it fails to capture certain concrete spatial features of insurgency and the similarity of insurgency's combat dynamics to the one of conventional warfare. Because of unidirectionality of lethal fire, as a physical property, by dispersing its force into multiple units both in space and in time the insurgent is reducing its exposure. This is done first by approaching the counterinsurgent from an angle not covered by its initially selected field of fire; it allows avoiding the lethal space established by the counterinsurgent at the initial time of combat, even after being

[65] I discussed the importance of the lines of attack earlier and will do it again below – they generate surprise by creating micro-instances in which one belligerent is totally exposed, while the other has a minimal level of exposure. Given the speed of bullets, the player attacked may not be able to re-direct his weapon against the attacking opponent before being killed.

spotted. Secondly, the insurgent explores the counterinsurgent's uncertainty about its location. This is well captured in the definition of exposure that I am applying, which indicates that the exposure of Blue is determined by both its presence in the lethal space of Red (physical access) and the awareness of Red about Blue's presence in that field of fire (temporal access). It implies that the exposure nature for both conventional and insurgency types of war is identical in nature, with a difference in its degree. So, we are dealing with two types of combat interaction that differ only through the intensity of one of its variables.

In fact, the exploration by the insurgent of both the surprise of physical access and the uncertainty about the temporal access is, arguably, at the heart of insurgency, as a model of warfare. The Russian pitiful experience in its first Chechen campaign is a good illustration of these principles. While popularly justified through the ignorance and insufficient technical skills of the Russian military, the repeated walking into Chechen ambushes of Russian units of various size, up to brigade level, provide the quasi-experiment that allows us to observe the effects of the exposure variable on militarily-asymmetric engagements, which is one of the key particularities of the insurgency model. During the start of the first Chechen campaign in December 1994, culminated with an assault on Grozny on 31 December, the Russian forces entered the city in march formation. This meant that tanks and other armored vehicles lacked infantry support and were vulnerable, unaware about their entering of the areas targeted by rebels. Moreover, they also had no maps of the city for units' commanders, while Russian officers used improper

maps during the planning phase and lacked advanced reconnaissance, which suggest they ignored the impact of exposure by being unaware of both its physical and temporal dimensions. On the top of it, the Russian superior force failed to blockade the city to prevent the resupplying of the rebels, among many other errors (Grau 1995, Lieven 1998, Miakinkov 2011), which is another failure to increase the rebels' exposure, by localizing them in targeted battlespaces. This superficial approach severely aggravated the effects of combat exposure, already asymmetrically higher for the counterinsurgent, as a rule. It prevented the Russian troops from diminishing their exposure to the extent it was possible in an urban environment, while failing to instantly increase the exposure of the opposing Chechen rebels. The result was an extremely high attrition rate for the Russian units and significantly reduced the attrition rate for the opposing Chechen rebels. Tanks and infantry armored vehicles, according to the quoted sources, were ambushed on the Grozny streets by whole units. This resulted in the Russian military loosing over a hundred different types of armored vehicles during that battle.

Another thing to keep in mind is that both in (b) and (c) the attacks of Blue's smaller units can take place at different points in time. By comparing and illustrating the two types of warfare in this way we can show that the insurgency warfare is in fact a subset of the conventional warfare, in terms of its combat attrition dynamics.

Perceiving these common features of the two types of warfare requires quite some level of abstraction from the particular contexts of each of them. It is, like in the modeling enterprise, when we need to explore a social phenomenon by stripping

it off of anything that is not essential for the understanding of its underlying dynamics of interest. Given that our unit of analysis is combat interaction, this simplification is justified. A very good interpretation of the differences between the two types of warfare is that the insurgency raises evasion over firepower because the insurgent can hardly win a direct contest of firepower. In response to an increasing evasion by his opponent, the counterinsurgent is forced to review and rethink its usual strategies, procedures, and even capabilities of target acquisition and targeting generally, as a way to diminish his emerging disadvantage.

Both from the figure 4.2 and from the combat interaction model introduced in chapter one, we could deduce a few more conclusions. The interaction terms (*rxy* and *bxy*) in the equation (17) suggests that by breaking its capabilities into smaller attacking groups, the insurgent is thus diminishing his exposure and casualties. This occurs since the usual size of casualties is a rate of the total size of the attacking group. When confronting an opponent more superior in firepower this approach is robbing the opponent of that important firepower advantage.

This dynamic process is more visible in case of conventional conflicts, due to combatants' larger sizes. For instance, at Somme the British casualties were so large because the attacking force that was sent into battle was so large. The larger the force involved in combat, the bigger in nominal size is the percentage of the casualties, determined by a fixed degree of exposure. Besides, by breaking the large force into smaller groups, they were distributed across multiple virtual battlespaces, requiring additional cognitive and logistic resources from the

defenders, which leads to a nonlinear decrease in the rate of casualties for each of the smaller groups.

This also explains the outcome of the French tactics at Somme. What the French side did, was to break down its attacking forces in smaller groups, thus virtually assigning them to more than one battlespace. By this they required separated weapon aiming, and attention of the defenders. It also required more time to aim and target the attackers, which contributed to the reduction of the period that the attackers were exposed to asymmetrically high degree of exposure. According to equation (33) the higher the difference in exposure of the two asymmetric belligerents,[66] the quicker their combat capabilities will approach to equality. Moreover, in line with equation (10) the French side, through its tactics, increased uncertainty about the number of virtual battlespaces occupied by its troops, while also contributing towards decreasing their exposure.

A similar dynamic is visible in various insurgent settings, when insurgents, for various reasons, decide to increase their forces' exposure. Some research pointed out they do so by trying to hold ground, switch from an insurgent to a conventional combat orientation, or engage in high stake offensives against the militarily stronger opponent (Krause 2009). Krause suggests these were important factors that set the stage for the decline or defeat of the Communist Party of Greece (Kommounistikó Kómma Elládas – KKE) insurgency, for the Salvadoran Farabundo

[66] The implication is that the weaker side has a smaller degree of exposure, either as conventional defender or as insurgent.

Marti National Liberation Front, and the Popular Front for the Liberation of the Occupied Arabian Gulf (PFLOAG), among others. For instance, Krause points out that in 1947 the KKE made efforts to defend territory in the Grammos-Visti region, both to attract international support for its self-declared government and to secure external support from Yugoslavia, which used the area as a logistic corridor (2009, 54). He also emphasized how PFLOAG, while trying to achieve a major military victory in 1972, attacked the town of Mirbat and suffered a painful defeat in which it lost almost half of its attackers (2009, 55). While these micro-examples reflect subsets of the more complex dynamics underlying these insurgency conflicts, they are just a few examples of the larger set of cases that emphasize the role of exposure in insurgency wars.

Now, let us consider some extra elements to the presented general combat attrition model. According to a large volume of existing literature, the assumption is that insurgent's reliance on population could bring them a few important benefits – contribute to the preservation of a low-level exposure, but also obtain logistical supplies and recruit new members. Each of these claims would ideally require a separate assessment, because they are context-specific. By this I mean that the statements about these important processes are often anecdotal and they frequently serve as untested assumptions in models backing theoretical and empirical claims.

For instance, the dependence on the population of insurgents' ability to maintain a low exposure is most famously reflected by Mao (1989, 93), who

compared the people with the water, and the insurgents with the fish who inhabit it. In many other cases, insurgents hid in mountainous areas (Chechnya, Afghanistan to name just a few), jungles (Vietnam, Colombia), or large areas of the desert (Iraq, North-Western Africa). We do not know to what extent these other insurgencies, except for rural and urban ones, critically depended on the local population in maintaining a low exposure level – there has been no thorough study of this, to my knowledge. Instead, these insurgents rather seem to be exploiting the cover of either extensive, or rough terrain, which is difficult to control or search by the counterinsurgent.

We could find strong support for this idea if we carefully examine the tactics and strategy of the Communist insurgency in China, under the leadership of Mao Tse-tung. This is very well illustrated by Boorman and Boorman (1966) who emphasized that the Communist insurgents adopted what the authors called a counter-encirclement campaign. The insurgents' defensive strategy was not to be tied to particular territory but to move around when either the Nationalists or the Japanese forces would come to engage them. The strategy contained three foundational elements, of which the first two were critical: (i) to preserve their base area (not a fixed geographic point); (ii) to preserve their own military force that defended the base area; and (iii) if possible, to annihilate all or part of the attacking force (Boorman and Boorman 1966, 178). The way this was implemented in practice was not to defend a particular area. Instead, when the opponent superior forces approached, the insurgents tried to avoid encirclement and to directly engage the

attacker. They moved around the countryside in a fluid manner, engaging sporadically the pursuing troops of the enemy that extended across the territory, breaking down its main force into smaller subsets and extending its communication and supply lines (Boorman and Boorman 1966, 180-181). This led to gradual wearing down of the attacker and eventual counterattack by insurgents when the former was weakened. In fact, this is the physical implementation of an exposure-maximization strategy, in line with the model intuition. However, this was in particular possible because the insurgents had extremely large terrain to explore. While the local population was important in providing the necessary supplies, the literature pays less attention to these characteristics of the possible battlespaces, such as size or difficult to control the area in which a counterinsurgent engages the insurgent.

Another corroborating example is the approach used by the United States military advisors and the Greek military do destroy the KKE. While conceptualized differently, it reflected similarities with the quadrillage tactics of the French counterinsurgents in Algeria, allowing territorial control. The Greek military have explored a set of consecutive though slow operations to locate, surround, and militarily degrade the KKE groups (Abbott 1994). The importance of controlling terrain was underscored by the ability of significant numbers of the Communist insurgents to flee to Albania or Yugoslavia, breaking through encirclements. The focus on methodical location and clearing of rebel-controlled areas underscores the essential character of exposure in insurgency wars.

The listed examples reflect both critical elements of the uncertainty behind the exposure concept – the uncertainty about the battlespace where the insurgent is located and about the partition of its forces across battlespaces.

In terms of material supplies and recruitments, practices vary across multiple cases. In some cases, insurgents benefit of supplies provision from loyal population; in others they had to forcefully take what they need, antagonizing the population, or purchase their supplies thus creating a pragmatic relationship with the civilians. The recruitment is also based on multiple contextual practices – in African conflicts it is not infrequent for the insurgents to capture children and train them into soldiers (Achivarina and Reich 2006), while, for instance in Colombia, others would join voluntarily due to the lack of economic alternatives (Bjørkhaug 2010). In Chechnya the recruitment was facilitated through teip (clan) affiliations (Galeotti 2002) or nationalistic feelings. This suggests caution is necessary when making assumptions on preventing recruitment to justify certain wider theoretical claims, made as part of the population-centric strategy of counterinsurgency. In particular, since lately the insurgents have proven richer than the population and used foreign-supplied money to pay locals for staging attacks against the U.S. troops (Kilcullen 2006) as an operational outsourcing strategy. Another research indicated that the resilience of an insurgent group may depend on its structure, its social network links, and control over resources, which could even lead to the group's depoliticization (Staniland 2012). Such insights could suggest different dynamics behind insurgencies' resilience, not necessarily dependent on the indigenous population. In fact, there is

a strong analytic insight that the focus should be on the supply of resources allowing the insurgent to compensate for its combat losses, and not on population, which is just a source for these supplies.

There are a number of empirical examples in support of this intuition. For instance, the KKE decline was significantly affected by its loss of support in weapons and other logistics coming from Yugoslavia, following the Stalin-Tito feud (Iatrides 1993, 215). A similar trait is reflected in the case of the Kurdish insurgency in Iraq, following Iraq success in improving its diplomatic relations with Iran, who was the main sponsor of the insurgents. In a matter of weeks, lacking sanctuary, arms and other supplies the Kurdish insurgency became dysfunctional (Krause 2009, 53).

Therefore, similarly to how this research insisted that combat dynamics deserves specific attention and investigation, to replace the black-box type of assumptions about how capabilities transform into victory, the claims about population support and recruitment require additional research attention as well. These claims, based on heuristics, are the Achilles's heel of many otherwise very solid and thoroughly conducted analyses. As examples, Berman et al. (2018) as well as Lyall and Wilson (2009) use anecdotes to introduce assumptions, which are critical elements justifying the existence of the data-generating process they claim to explain.[67] One could argue that the former study triggers somewhat less severe issues, given the popular style of the book presenting scientific research on the topic. The latter analysis, however, attempted to justify the claimed mechanism of

[67] This references' list can be greatly extended, though this is beyond the scope of this research.

“information starvation” referring to interviews of U.S. officers in Iraq (2009, 97). Provided these statements were part of the official U.S. military’s doctrine at that time, the interviewed officers would not have been able to describe the situation otherwise. And even if they challenged the official policy, they are only good observers, who are not trained in the scientific method. It would require a more rigorous examination of the case to scrutinize the validity of their claims, which could at the time, mostly, be justified as hypotheses. In fact, Smith and Toronto (2010) provide a detailed and well-argued account of counterinsurgency microdynamics, which questioned Lyall and Wilson assumptions on “information starvation”. This is another example how important is to build our assumptions on a sold understanding of the microdynamics affecting the phenomenon of interest. It is this type of practices that recently drew criticism from prominent qualitative researchers (Mearsheimer and Walt 2013) and provide ammunition to the critics of quantitative methods in social sciences.

For now, however, let us accept these assumptions at face value, and explore them through the combat dynamics model introduced earlier, to assess how they can affect the rate of attrition. An important parameter that was not yet treated directly is the support that the insurgent receives both from the population and possible foreign sources. In fact, drawing support from the population along with being able to hide among the population, as an exposure-diminishing strategy, has traditionally been viewed as the secret of insurgencies’ survivability and potency against militarily stronger counterinsurgent. The role of these continuously incoming

supplies is to compensate for the losses of insurgent forces inflicted upon them by the counterinsurgent. Without these supplies the insurgent would eventually lose all of its fighters to attrition and vanish.

In the above model (27) we can imagine the supplies entering through the constant of integration c_y, which is an initial condition of the system, representing the insurgent troops level at the beginning of the military confrontation. For simplicity, we can visualize that continuous supply of troops and resources flow, as just transforming into a larger value of c_y. The duration of time, during which this number of fighters will survive will still depend upon the size of the eigenvalue, which is the exposure rate. In fact, if we give the insurgent a larger number of troops but, in contrast, make his exposure higher than the one of counterinsurgent's troops, the insurgent is going to be defeated, as illustrated through a range of graphs in the second chapter (fig. 2.4-2.11). Exposure, therefore, remans the essential variable in insurgency as well, holding primacy over capabilities and firepower.

Let us further illustrate this logic with another visual aid. Figure 4.3 offers a complete description of the combat dynamics determining the attrition of the insurgent.[68] From the dynamic system captured in the equation (16) we know that there are two drivers that lead to the insurgent's attrition – the one dependent on its exposure, marked here as $-\mathcal{E}_{rp}R$, and the interaction term reflecting the direct combat engagement with the counterinsurgent, $-bBR$.

[68] The focus is only on the insurgent, because it is his asymmetric exploitation of exposure that makes insurgency a conceptual subset of conventional warfare, based on attrition dynamics. Doing the same for counterinsurgent does not help advancing my argument, which is why I am ignoring it.

Even though this would be a simplification, we could assume the exposure-dependent term to be linked to the safe heaven that population might offer the insurgent, including the private information it may have about the location of the insurgent.

However, the system of dynamic equations (16) did not implicitly model that the insurgent may be additionally getting some material assistance from the population, as well as recruits, $+R_p$. For the understanding of general combat dynamics this is not essential.

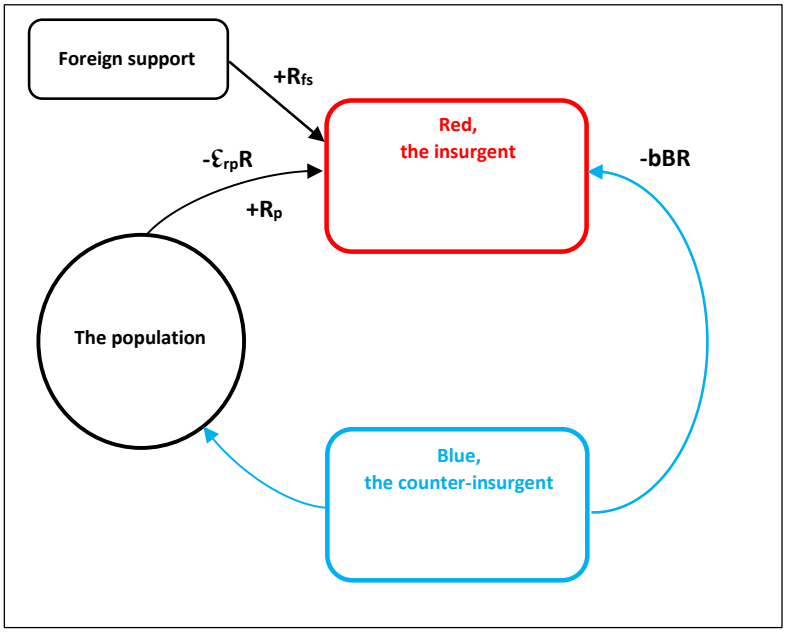


Fig. 4.3: Schematics of the insurgent's attrition process and variables affecting it

Here, however, it is useful to consider it for a better illustration of insurgency attrition dynamics. In particular, this is useful to explore as so many analysts are making

strong claims and assumptions about the role of this type of support for the survival of insurgents. In this case, the $+R_p$ quantity operates by compensating the losses accumulated by the insurgent through his combat interaction with the counterinsurgent, $-bBR$.

This is an important remark, because the insurgent has finite resources, which means they will get eventually depleted in attrition (combat) interactions with the counterinsurgent, if not compensated. In fact, as Fearon (2004) shows it, if insurgents identify a source of considerable revenue - either exploring local natural resources, smuggling schemes – they are likely to be able to survive, resisting the stronger player for extensive time even without population support. Foreign assistance in cash, if provided directly to the insurgency leadership would very likely produce a similar impact. This funding allows offsetting the attrition generated by the counterinsurgent. The insurgency will flourish or survive for a longer time if either $(R_p + R_{fs}) - \mathcal{E}_{rp}R \gg bBR$[69] meaning the left side is significantly larger than the right side of the equation, or respectively, if the two sides are approaching in value. As the model and the illustration above suggested, under these conditions only an elevated level of attrition against insurgent could allow the counterinsurgent to reach the threshold of effective violence (Zhukov 2014).

It should be mentioned, though, that a population-centric strategy could increase the insurgent exposure significantly, under the conditions when the

[69] I included the $-\mathcal{E}_{rp}R$ quantity on the left side of the equation, to show that the insurgent exposure rate can be decreased as a result of positive population-insurgent interaction, which will result in a decreased attrition (though still positive) of the insurgent's forces.

insurgent's only refuge is provided by the population and it does not have other hiding places except dissipating among the population. It is though unrealistic to assume that a population-centric approach could prevent supplies and recruits going to the insurgent, as these can be acquired through coercion or other negative incentives, operating either as a stationary or roving "bandit". In particular, it is hardly possible for a foreign counterinsurgent to completely insulate the insurgent from the assistance of true local sympathizers. It is under this conditions that states tried to physically separate the population. As French (2012) quoted Sir Richard Turnbull, the Chief Secretary to the government of Kenya between 1955 and 1958 "… how could we as strangers win the hearts and minds of the Kikuyu in opposition to their brothers and cousins who were in the forests?" This means, the $+R_p$ is going to be most of the times positive, especially when the counterinsurgent is either a foreign actor or is directly assisted by a foreign actor. This would strongly suggest that population-centric strategies would not increase much the attrition of the insurgent if the latter is getting most of its assistance from abroad, as implied by Toft and Zhukov (2015).

Some other analysis, while providing a convincing criticism of the population-centric approach, argued that the key to counterinsurgency success is achieving the control of the population instead of winning its "hearts and minds".[70] This would not significantly undermine the insurgent attrition rate, if the latter also receives $+R_{fs}$

[70] That control implies a combination of three tactics: physical measures, cooption and 'divide and rule' efforts. The physical measures, according to Cohen (2014a), include building walls, controlling resources and forced resettlement, among others. The cooptation refers to coopting local elites or elements of the insurgency.

support from a foreign actor. In fact, there is a voluminous research indicating that foreign sponsorship of insurgencies is of critical importance for their survival (Byman et al. 2001), supporting the theoretical framework presented in relation to figure 4.3.

We should also consider the possibility that the supply of these resources by the population to the insurgent is endogenous to the level of violence inflicted by both combatants, including the counterinsurgent. Or, as frequent anecdotal evidence brought by academic researchers suggests, in words of a U.S. military officer "we suspected that for every Iraqi killed or injured by U.S. forces, we were essentially creating more new insurgents" (Berman et al. 2018, 7). This logic would suggest that the quantity of losses inflicted by the counterinsurgent would be $\Delta R(-bB - \mathcal{E}_{rp})$, where $\Delta R = R_0 + R_p + R_{fs}$. This illustration again confirms that ΔR is just a reservoir, from which insurgent draws capabilities and which the counterinsurgent strives to empty. This can happen through two negative feedback loops, one representing the interaction term and another one the general attrition term – both of them diminish the insurgent's reservoir ΔR. However, according to the model, the counterinsurgent still has the upper hand as long as it manages to deplete more resources than the insurgents attracts from its local and foreign supporters. These supplies are finite. The insurgent is in a very vulnerable position, in fact. The counterinsurgent can deplete more resources by increasing the level of attrition, through intense violence. There is no evidence supporting the claim that for every insurgent killed two more join the insurgency, and there is no logical reason to accept this claim. However, the mathematical modeling would suggest that even

if this is the case, the counterinsurgent still has the upper hand, when keeping a high level of violence, because it is impossible to maintain for long the 2:1 replacement.

An intuition into this is offered by O'Loughlin et al. (2010) in their analysis of violent incidents and data on casualties for the insurgents in Afghanistan, based on the analysis of the WikiLeaks Afghanistan War Logs.

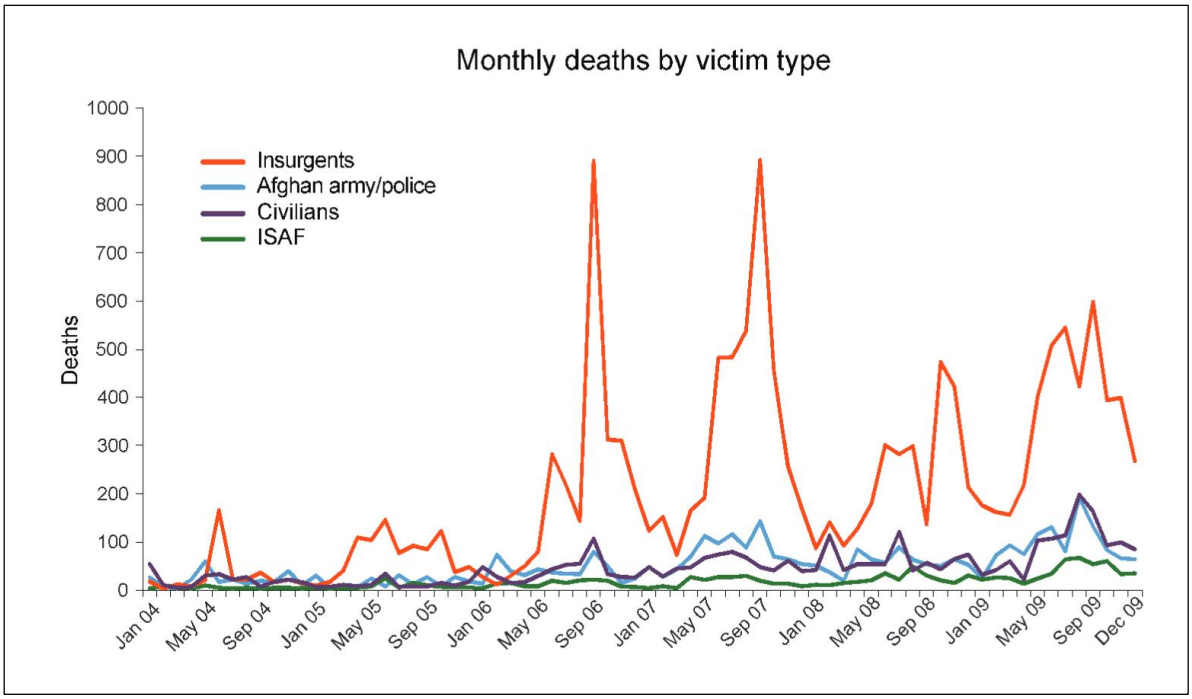


Fig. 4.4: Monthly deaths by victim, Afghanistan 2004-2009[71]

The image suggests that the violence occurs in spikes. There is operational logic behind it – you need to logistically and operationally plan the attacks, recruit, train, and mobilize people, buy arms, ammunition, explosion material and other

[71] Compiled by O'Loughlin et al. (2010, 479) based on WikiLeaks data.

supplies. We can notice approximatively one spike of large violence on the side of insurgents per year. Provided that the number of insurgents is limited, a combat tempo like this would seriously undermine the insurgents' capacity to continue exist as a relatively strong military force, able to exercise control over some regions and extract resources. Their reservoir is not bottomless. O'Loughlin et al. (2010, 476) suggested the insurgent lost during 2004-2009 over 15,000 killed, over 1,700 wounded and some 7,200 detained. While the counterinsurgent is able to sustain this combat tempo, considering it had significantly less deaths,[72] the insurgent would unlikely be able to maintain this trend for long.

In fact, some sources (Jones and Johnston 2012) pointed out that the Taliban deployed historically from 20,000 to 40,000 forces, the most, suggesting that in the time-span reflected above, the Taliban potentially suffered some 50 percent attrition of its capabilities, which is a very serious loss for an insurgent.

Before accepting heuristic claims about insurgents recruiting following indiscriminate violence carried by the counterinsurgent, we should ask ourselves what are the incentives of the population to join the insurgency and fight against the counterinsurgent? In fact, the existing rational choice models would suggest that the higher the level of violence maintained by the counterinsurgent, the less incentives the civilians would have joining the fight. It is exactly the low level of violence generated by the counterinsurgent that would create incentives for the civilians to

[72] ISAF lost, during the same period 1,140 killed, while the Afghan National Security Forces (ANSF) – 3,754 killed. The number of wounded is much higher – over 7,000 for ISAF and over 8,000 for ANSF.

support the insurgent, given the high opportunity cost. This is especially the case when the citizens are coerced to do so by the insurgent, or if they are attracted by economic benefits, like per-attack-payment.

These points find support in various examinations of combatant recruitment in insurgencies. For instance, Humphreys and Weinstein (2008, 447-449) find evidence that in the case of the lengthy civil war in Sierra Leone, the recruits for various insurgent factions were significantly attracted by material rewards, joined after being coerced, or considered being part of the insurgency was safer then outside it. Economic considerations behind population support for insurgents have been influential in understanding the lengthy survival of insurgencies in Colombia and Peru. Thus, the protection offered by Shining Path to coca-producers in Peru and the FARC protection to farmers of coca and poppy in Colombia, along with the governmental efforts for drug-eradication, have reportedly led to higher popular support for insurgencies (Felbab-Brown 2005).

This evidence would strongly suggest that the decision of the population of either supporting or not the insurgents, including by joining them, is based on a much more complex strategic consideration. Depending on the initial conditions of the particular combat environment, the population may choose to support the insurgent either to protect its gains, or to minimize its losses. A more thorough examination of this recruitment mechanism is necessary, as micro-evidence from various conflicts tends to challenge the “two insurgents for one killed” heuristic claim.

A good illustration of this point is the Lyall et al. (2013) research on public attitudes of population in response to violence inflicted by the Taliban on one hand, or the International Security Assistance Force (ISAF) on the other. While the surveyed population, representing 204 villages in five Pashtun-dominated provinces in Afghanistan, was likely to increase its support of Taliban in response to ISAF violence, it was unlikely to increase its support for ISAF in reaction to Taliban violence. Since Taliban generated a higher level of violence against the population than ISAF, generating higher costs for the population, the quoted research would question both the premises and the conclusions of the population-centric view on insurgency war and would indirectly support the general logic of the proposed exposure-attrition model.

Unless the counterinsurgent is able to effectively control the area, being able to insulate the population from the insurgent punishment, a population-centric logic lacks a conceptual mechanism supported by evidence. To control effectively the area, the counterinsurgent needs to significantly degrade the capabilities of the insurgent, as suggested earlier. Otherwise, the population's incentive remains to diminish its losses, which are higher in the interaction with the insurgent. In fact, Lyall et al. (2015) show convincingly that the Pashtun population was unlikely to collaborate with the ISAF due to the high effectiveness of Taliban in identifying and punishing the ISAF local informers. The effort to reduce indiscriminate violence affecting the population is a necessary goal for a democratic counterinsurgent, but not for the reasons the population-centric literature on insurgency claims. The

exposure model has more promise in explaining the dynamics of insurgency wars and its reasons for success or failure.

Exposure in insurgency warfare

Based on the insights offered in figure 4.3 we can examine the essential factors that either speed up or slow down the insurgent forces' attrition. Applying brute force against the insurgent or the enemy-centric strategy is an essential condition for achieving that end. As the model suggests, there is no alternative in counterinsurgency for military force. When the insurgency is strong, police is unable to quell it alone (Gentry 2010), in particular since it is not equipped and designed for this task. A number of historical cases, some of which were referred earlier in the text, offered strong evidence in that regard.

In order to win the counterinsurgent would need to increase the speed of attrition of the insurgent troops up to a level, when these cannot be compensated by the support that comes from either the local population or the foreign sponsor. Alternatively, the counterinsurgent would have to interdict somehow the supply of reinforcements, both in material and in recruits, while maintaining a positive troops' attrition rate for the opponent. The essence is to be able to affect both the input and output sides of the insurgent's resources, instantly.

Finally, the counterinsurgent should not forget to optimize its strategic interaction with the insurgent in terms of increasing the opponent's exposure while

decreasing his own. It is very often the case that this second part is ignored both in academic research and in policymaking. Due to its modus operandi, this is what maintains the exposure of the counterinsurgent as such high levels – its stationary military bases, its offices in cities, its patrols and teams interacting with the population, its detached and open observation posts in the country-side, its logistical supply lines, among other things. Some of the seek-and-destroy techniques referred throughout the chapter address partially this inherent problem for the counterinsurgent. More needs to be done though and this is a critical challenge for military operations researchers and political scientists to tackle, with potential high added value to counterinsurgency research.

There is more than one way to reach that goal. In game theoretic parlance the counterinsurgent could pursue mixed strategies in terms of how much to focus on the attrition of insurgent troops directly, and how much of its resources to direct towards diminishing the support that the population, or a foreign actor, lend to the insurgent. In line with a population-centric strategy the counterinsurgent will sacrifice the efforts to put attrition pressure on the insurgent in favor of the efforts to better control the interaction space between the insurgent and the population in the theater of operations. An enemy-centric approach will lead the counterinsurgent to degrade the capabilities of the insurgent ignoring the collateral damage.

Let us examine, for illustration purposes the Russian experience in the second Chechen campaign (1999-2000), following the humiliating defeat the Russian military suffered in its first military campaign (1994-1995). Miakinkov (2011)

analyzed the Russian counterinsurgency practice and came with an interpretation of it. He contrasted it to the US counterinsurgency approach, that requires a “flexible, adaptive force”, which can match the tactical agility of insurgents and aims at winning the “hearts and minds” of the indigenous population. The Russian set of tools were different. Miakinkov develops a model of Russian counterinsurgency, arguing it included overwhelming superiority of force, and cultivating the support for the military operation among its domestic audience. Because of different perception by the home public of the indigenous population, the Russians prioritized the support of home public, not the support of indigenous population. While, perhaps, generating identical dynamics in the late US counterinsurgency cases, it generated a different dynamic in the Russian case.

In Miakinkov view, the second element of the Russian strategy ensured that the first one is explored without restraint. Exposed to an intensive propaganda campaign, the Russian public was less critical of the harsh methods the Russian military used in Chechnya. As he puts it:

> A successful counter-insurgency campaign is contingent upon winning the support of the public at home, so that the conventional armed forces can deploy an overwhelming superiority of force to contain and then neutralize insurgency threats. In other words, this work submits that insurgency can be suppressed and contained through force alone, provided that public support at home is secured.

According to Miakinkov (2011, 668-669), the unrestrained violence perpetrated by the Russian military during the second Chechen campaign produced another operational consequence. The indigenous, Chechen people, started to grow tired of the conflict. Miakinkov indicates that by 2000, the "majority of the Chechen population was supportive of actions aimed at restoring order". Following the failure of the insurgent to show success the hope of the population was lost and it was not any longer ready to pay the costs of liberation from Russia.[73] This was a consequence of how Russian military decided to increase the exposure of their Chechen opponents. Russian military would surround the area where it believed the Chechen militants were stationed and would bombard that area with heavy artillery and aviation for a few days. Following this, it would send troops into the area to remove the remaining opposition. This dynamic of population conflict fatigue is less likely to be observed in areas where conflict is the normalcy, like in Afghanistan.[74]

Thus, while this Russian tactic was frequently viewed by the Western analysts as a punishment strategy, it was far more pragmatic. It instead aimed at preparing its target and soften it up, for the follow-up engagement by the Russian strike groups. This would allow the Russian planners to tremendously reduce own casualties, comparing with the first Chechen campaign. For instance, after reaching the capital city of the rebel region, Grozny, in mid-October, the Russian military

[73] One could argue this can reflect a strategic goal of the Russian military, to use indiscriminate violence in order to reduce the Chechen civilians' support for the insurgency, as the fighting was destroying their way of life.
[74] If, due to lack of economic opportunities, rural population views occasional participation in skirmishes and bomb attacks of the insurgent war lords as an attractive economic activity, conflict fatigue is less likely to affect these groups.

bombed it until December 1999 (Oliker 2001, 42-43). The city was also encircled, to reduce the chances for the few thousand rebels that defended the city to escape it or get resupplied – another tactical approach of the Russian planners to increase the Chechen militants' exposure. Even though in practice the city blockade was porous (Oliker, 2001), it did create uncertainty among the rebels about facing Russian patrols and getting intercepted, increasing the insurgent costs.

The Russian military applied these tactics of massive artillery and aviation bombardment to other cities. On 7 March it approached and surrounded the city of Komsomolskaya, where some 1,000 Chechen rebels were believed to have been taken defensive positions (Miakinkov 2011, 670). This heavy-handed approach has inflicted severe casualties among the Chechen militants, so that they withdrew to mountains. Here, again, it is illustrative to see the tactics employed by the Russian military to increase the exposure of the rebels. The Russian military set up patrols and check points on the main transportation arteries and roads that connected the mountains with the inhabited areas, such as cities and villages. The rationale behind this approach was that the rebels would eventually have to search for re-supplies. This was the Russian interdiction and containment strategy. By separating the rebels from the inhabited area, they also made it more difficult for the rebels to conduct recruitment (Miakinkov 2011, 672).

The tactical choice of using heavy bombardments against the rebels' positions or their believed location, instead of storming them with infantry like in the first Chechen campaign are reflecting the Russian military planners' solution to

diminish their troops exposure against an agile insurgent. It is imperative to notice that unlike in the previous campaign the Russian military invested significant effort into reconnaissance and observation of the rebels prior to the operations. This was done through insertion of small airborne groups, special forces operations, aerial observation and work with local informants (Thomas 2000, Oliker 2001). Thus, my model's exposure variable provides the suitable theoretical framework necessary to explain alternative interpretations of indiscriminate violence (as substitute for intelligence). In addition, the Russian bombardments were also attempting to maximize the attrition impact on insurgent's troops. The bombardments were extremely intense – they even used fuel-air explosives and tactical missiles. The former were exploited to target the rebels hiding in the basements and also had significant psychological effect (Thomas 2000). The Russian military tried to elevate the Chechen rebels' exposure by saturating a wider area with massive firepower, targeting even the urban environment in which it believed the insurgent to be hiding. This justifies conceptualizing exposure in spatial terms of being *targeted or accidental*. In terms of the terminology and concepts from the first chapter, the counterinsurgent was thus trying to aggregate a few adjacent battlespaces to prevent the dispersion of his capabilities and the erosion of this advantage, while attempting to maintain the speed of attrition of his opponent. This theoretical interpretation confirms the conceptual intuition of the theory that this work introduces.

If, the US and Western approaches seem to be in courting the population, to separate it from the insurgent, the Russians strategy is different because they did not trust this was possible. Instead, they coerced the insurgents to flee the inhabited areas, forcefully separating them from the population. This allowed the Russian military to reduce the input (supply) side to the rebel attrition dynamics, while continuously targeting its output side, which was reflected in its combat capabilities. It forced the rebels to lose more personnel and resources than they could replenish.

A similar approach, perhaps accidentally, was implemented by the French military more recently during the Operation Serval. The operation was conducted by the French military against the Tuareg separatists and Islamist extremists in Mali in 2013-2014. Following the risk of rebels and Islamists with links to Al Qaeda in the Islamic Maghreb to take over the Mali's capital city, the French send an expeditionary force of a few thousand troops at the invitation of the provisory Mali government (Wing 2016). Observers reported that French operation, combining tactical airpower and special forces, managed to stop the rebels' offensive in a matter of days. French forces took control over important cities in the northern part of the country. The rebels reportedly withdrew to the mountains along the country's north-eastern border with Algeria and Niger (Boeke and Schuurman 2015). Having secured military support from a number of African countries, the French forces engaged the rebels in combat over the next few months. This resulted in some 30 percent of the 2,000-strong Islamist militants force being killed and over 400 being

taken prisoners. The French casualties were insignificant, not exceeding a few dozen.

Another effect of the brute force approach used by the Russian military against the inhabited areas were that the villages' authoritative informal leadership – the community elders – realized that assisting the rebels would most likely result in the bombardment and destruction of the village. The outcome logic was not that of an indiscriminate violence, as interpreted by Western observers, but of a "collective punishment". In that case, the mistake of a member of the community might trigger the punishment of the whole community, creating incentives for the community to discipline its members and prevent their non-cooperation with the Russian military. A mechanism, showing similarity in its dynamics and outcome, reflecting a logic of in-group policing, was discussed and described by Fearon and Laitin (1996) in a different context.

This Russian experience during its second Chechen campaign provides a few other insights. Existing counterinsurgency research suggests that some useful metric in planning operations is either the force-to-population ratio or force-to-force ratio.[75] These metrics have even guided the US military official policy towards counterinsurgency warfare. However, as the Russian experience suggests, the force sufficiency is effectively guided by combat tactics used and geography. In mountainous areas, where the movement and supply are possible through the transportation arteries, which create a graph or network, controlling effectively that

[75] For a detailed discussion of both ratios see Watts et al. (2014).

movement can be done by optimally placing check-points at critical vertices, and patrolling specific edges of the graph.[76] This logic is not new as the highly-effective, from a military point of view, French counterinsurgency operation in Algeria used a procedure called quadrillage to separate the combat environment and establish a more effective control of the French military over it (Melnik 1967). This is another example of how a combatant can reduce the opponent's evasion and thus increase its exposure.

The quoted differences in designing methods to control and restrict the input part of the attrition – the supply in personnel and material coming to the insurgent from the population – emphasize another factor that may significantly affect the insurgency outcome. Applying suboptimal control over these inputs ($+R_p$ and $+R_{fp}$, fig.4.3), will undermine the counterinsurgent's effort to effectively degrade the insurgent capabilities through combat attrition. It is the counterinsurgent's ability to keep the level of insurgent's combat attrition higher than the volume of incoming resources that will eventually lead to the defeat of the latter. Understanding these fundamental dynamics is helpful for accurately perceiving the effect of various factors (or other input variables) on success in insurgency or asymmetric warfare.

For instance, Mack (1975) argues that an actor's relative resolve or interest would explain success or failure in asymmetric conflicts. He claimed that the actor

[76] The Russian experience, thus, is suggesting that these studies attempting to determine optimal force ratios might have a fundamental flaw in their inference. They examine multiple counterinsurgency cases by applying some sort of averaging rule, assuming the data-generating process is similar across the majority of examined cases. In fact, it might be a better idea to explore the internal dynamics of the cases, separately, and examine the involved spatial considerations in attempting to establish an algorithm of optimal area control by a military force.

with the highest resolve is likely to win, regardless of the material or power resources under control. However, the proposed model of combat attrition dynamic does not support this logic. Similar to the discussion of the effect of morale on conventional war success, the resolve in Mack's model can enter the combat dynamics through a contest success function, like the one described by Hirshleifer (2000). Mack's resolve is affecting the combat attrition and attrition speed through the human combatant, or war planner - specifically through their decision to withdraw from combat at a later stage, after having suffered higher casualties.[77] However, as shown in the previous chapter, it is the exposure degree and the resulting troops attrition that determines who remains victorious on the battlefield. In fact, as explained earlier, resolve can lead to higher troops' attrition, by increasing the time of troops' exposure to enemy's lethal fire.

Due to the asymmetry of combat capabilities the effect of exposure and attrition is even more impactful in the case of insurgency war. In fact, if it wants to fight another battle tomorrow the insurgent will have to constantly withdraw from its battles, make efforts to reduce its combat attrition rates or keep them at limited levels. What Mack labels as resolve, a clear definition of which he and other scholars working on this analytic tool do not usually provide, is more likely to be reflective of the valuation attached to the good contested in the conflict. This is one of the key reasons why the Russian counterinsurgency strategy in Chechnya focused not on

[77] I sustain that the effect of Mack's resolve on combat is similar to the one that I investigated in the case of morale in the second chapter for the reasons invoked there.

winning the support of the indigenous population, but on winning the support of the general Russian voter. If the support of the latter for war is low, this would increase significantly the costs of that war; it will add to them the political costs, which are much more salient for the Russian policymakers than the military-related costs. Therefore, an essential element of the Russian counterinsurgency campaign in Chechnya during 1999-2000 was the state control over the media coverage of conflict (Miakinkov 2011, 674).

To describe the richness of policy choices in the operationalization and manipulation of the degree of opponent's exposure, we shall still examine one more of the Russia's counterinsurgency strategies used in Chechnya. To reduce the exposure of its troops and increase the exposure of the rebels, the Russian government has co-opted one of the influential Chechen clans to become the poster-boy for counterinsurgency in Chechnya. As a result of the deal, Putin selected the Kadyrov clan to govern Chechnya in exchange for its contribution to the quelling of the Chechen secessionist rebellion. Some authors referred to this approach as the "Chechenisation" of the conflict (Russell 2011).

The chosen model offered a few advantages from the point of view of a counterinsurgent, which generated an impact in diminishing the Russian military and political exposure. This, consequently, decreased the costs that the Russian leadership was ready to pay for its ability to suppress the rebellion. First of all, a large part of the combat was carried on by the Kadyrov's clan militants, many of whom were part of the anti-Russian rebellion in the past and were intimately familiar

with the insurgent tactics (Šmíd and Mareš 2015). This, arguably, allowed them to be more effective at increasing the rebels combat exposure, and subsequently the speed of their attrition. That consideration, though based not on acquired skills in combat, or higher level of applied violence, but on the co-ethnicity, was explored by Lyall (2010).

It has to be noticed that another advantage of the Kadyrov's militants was that they could engage in high-intensity violence against the militants, including against their families, as a way to put pressure on insurgents (Dannreuther and March 2008).[78] The Chechen militants were not the only ones to explore this family-targeted coercion. The groups that conducted this kind of operations included both Kadyrov's militants and the Russian special forces or intelligence agencies' personnel. These arrangements allowed for the political insulation of the Russian government from criticism and the resulting penalization by both international community and domestic audience. It offered the Russian government the ability for denial and permitted avoiding attribution.

To give it another interpretation, the fight with insurgents was outsourced to a local interest group, whose members would pay the cost for suppressing the insurgency, in exchange for the permission by the Kremlin to govern the region as they wished. In terms of the dynamical model I proposed the Russian side managed

[78] Some authors interpret this higher potential of local actors for violence, as being a more effective co-ethnic trait from an informational endowment standpoint. Based on the examination of these articles, we can see the interpretation was constructed heuristically, instead of examining the microdynamics of the data-generated process of interest. What these authors named more effective co-ethnics reflects just higher-intensity violence. See Lyall (2010) for an example.

to drastically reduce its exposure, slowing down the attrition of their forces and resources, including in the very important political dimension.

This effect – the increasing of the opponent's combat exposure along with the reducing of one's own exposure - is a critical factor for success in any war, including in an insurgency conflict. Most literature on counterinsurgency, to my knowledge, pays very little attention to this important intuition that my model allowed to derive. However, when instrumentalized it works not only in the case of authoritarian countries. The well-known case of "Anbar Awakening", when some Sunni tribes in Iraq decided to go against the al-Qaida in Iraq in 2006-2007 is a very good illustration of the universality of outsourcing the counterinsurgency combat effort. While the literature debates whether the "awakening" could have been successful with or without the surge in troops that the US government enacted at the time, it is largely accepted that no success would have been possible without the enlisting of the local interest groups in the counterinsurgency combat actions (Biddle et al. 2012).

The way of intelligence

A key claim of the proponents of the population-centric paradigm is that by responding to population's needs and gaining its trust, the counterinsurgent would start getting valuable intelligence on the location of the insurgent. Consequently, acting upon this intelligence, the counterinsurgent would be able to target the

insurgent and defeat it through attrition. The targeting would become more effective, because this intelligence on the insurgent's location would increase its forces' exposure. This very attractive story is built on a number of questionable assumptions.

Analysts point out to lack of thorough research in that domain (Krause 1996). Popplewell (1995, 350) quoted in Duyvesteyn (2011) even forcefully argues that "cases when intelligence alone was responsible for suppressing insurgencies are wholly absent from all current works on counter-insurgency". To further clarify this point, we would need to understand the underlying dynamics of the counterinsurgent-population interaction in terms of intelligence sharing, by building a dynamic model similar to the one used to understand combat attrition dynamics. However, this effort is beyond the scope of this analysis. Instead, we can employ the intuition of analytic narrative method (Bates et al. 1998, Levi and Weingast 2016) to try to understand some important traits of this relationship.

Let us start from the needs of the insurgents, which include human and material support, most of which cannot be supplied by local population (Byman et al. 2001, 84). These are arms and ammunition, financial resources, medicine, fuel, food supplies, training, organizational aid and intelligence. The local population could offer ad-hoc and reduced food supply, occasional safe heaven, and some intelligence, which is time-sensitive. The insurgents, as a rule, do not have the capabilities to control the population, as they only can aspire to make it either friendly or neutral. Exceptions are cases when the insurgents, by exploring weak

state capacity can establish quasi-administrative control over a region. This scenario is outside of the scope of this research; moreover, it is usually addressed with conventional attacks by the central authorities after which the classical insurgent-counterinsurgent interaction begins.

Therefore, the intelligence received by the population on the insurgent would be generally of a situational nature and time-sensitive as well. It will be based on the ad-hoc visits by the insurgents into the inhabited areas, when insurgents are based in difficult-to-access terrain such as mountains or jungles. In the case of urban insurgencies, a large part of them will be hiding in inhabited areas, being good citizens by day and guerilla by night. In these cases it will be difficult for the counterinsurgent to receive intelligence from the population, which will tend to be loyal to the insurgents. This has already been corroborated by the reference to Richard Turnbull, who based on his experience in Kenya questioned the ability of a "stranger" to earn the "hearts and minds" of the locals, given the insurgents were locals as well (French 2012).

The key point is that the intelligence-provision by the population to diminishing the exposure of the insurgent is not critical and cannot be justified as an implicit assumption of the population-centric strategy. This is not to say that attempts to diminish the population's grievances towards the counterinsurgent are unimportant and to be avoided – it is only suggesting they are unlikely to bring critical changes to operational conditions. And it suggests that alternative operational methods exist to increase the insurgent's exposure in no less efficient ways.

Secondly, the population-centric claim on intelligence provision reveals a complete misunderstanding of the intelligence-collection process in insurgency-type of wars in particular and in combat, generally. As one of many possible examples, a widely-quoted political science research that made related assumptions and built the argument on it, is Lyall and Wilson (2009). The authors posed that because modern military uses armored vehicles in counterinsurgency operations, this diminish their ability to interact with the locals and collect useful intelligence. As it will be developed further in the text, this suggest a lack of understanding of the intelligence collection process in counterinsurgencies as well as of the related historical practices. Smith and Toronto (2010) offered some empirical argument that undermined the criticism against mechanization in counterinsurgency. However, they did not sufficiently address the questionable assumption that is foundational for Lyall and Wilson argument on intelligence collection practices.

In fact, most intelligence collection is based on exploring local sources. As insurgencies usually emerge in the context of a civil conflict, there are population groups that either support or oppose the insurgents. Moreover, intelligence collection from population loyal to insurgents is done practically in a covert way, with deployment of observation posts, UAV observation, and the informant networks of local police and of the counterinsurgent's own intelligence units. These would suggest that the population-centric hypothesis, which claim earning the sympathy of the locals is the solution to the conflict, are at least unfamiliar with the operational constraints on the ground. This is why many practitioners and policy researchers

tend to talk about population control instead, where measures are taken to literally separate the population from the insurgents though more drastic measures like displacements or the distributions of photo I.D. cards among the population in tandem with networks of check-point controls (Sepp 2005, 10).

Third, the existing research that to an extent or another examined the information issue in insurgencies does not lend support to these population-centric assumptions. For instance, Hack (1999) examines the British counterinsurgent efforts in Malaya and indicates that intelligence was gathered through civilian police, through surrendered insurgency members and especially through population control. In fact, he places the highest weight on population control, suggesting the British did not rely on the local population to provide them with intelligence. An example Hack offers involved starving the rebels, interdicting food supplies, and then trying to apprehend them when they would come to collect food. Recent lessons learned from Iraq encouraged the U.S. intelligence community to rely on what they refer to as ethnographic intelligence, meaning the cultural awareness and understanding of the operational environment through specifically trained and deployed personnel (Renzi 2006). As Teamey and Sweet (2006) described, the intelligence in counterinsurgency involves all personnel, including patrols, doctors, engineers, Civil Affairs team members, among others, and is similar to law-enforcement information collection. Frequently, the patrol debriefs may be a primary form of intelligence gathering in many areas, collecting observed changes in tactical and operational environment, rather than talks with the population. However, it is

just one of many threads of intelligence collection that emerge together to complete the situational awareness of the military planners and policymakers.

An important and underexplored challenge of relying on intelligence from the population caught in between the insurgent and counterinsurgent, is the reliability of this intelligence. There is a complex social and political dynamics on the ground, ongoing in parallel with the insurgent war. This creates significant incentive on the side of population to mislead the counterinsurgent or the insurgent, to solve its local problems. Kalyvas (2006, 183) provides diverse and rich evidence in this regard, revealing the risks of entrapment by relying on population intelligence. He describes how a villager in Afghanistan could falsely denounce a competitor from another village to the American forces, in the hope that he would be arrested; similar examples from the conflicts in Guatemala, Vietnam, post-civil war Russia, WWII-time USSR, Philippines or Iraq emphasized this disinformation hazard.

This is why the best and more accurate intelligence collection is obtained by aggregation from different sources. As military intelligence personnel reveal, it is not infrequent for intelligence to be collected through specific triggering actions – cordoning a neighborhood and conducted door-to-door checks would trigger reactions among potential information sources that may try to escape the area and are observed by unmanned aerial vehicles or cordon-and-search forces (Zeytoonian et al. 2006). Practitioners also emphasize that while human intelligence is very important, other tools, such as signal intelligence, image intelligence or measurement and signature intelligence continue to have major value. Therefore, if

mechanized patrols are able to interact less with the locals, due to their vehicles, it does not mean that the unit is deprived of necessary intelligence and situational awareness, which is achieved through many more ways and procedures.

To illustrate this let us consider the Lyall and Wilson (2009) well-known article claiming mechanization is undermining the efficiency of counterintelligence. Ironically, their illustrative comparison between the light 101st Air Assault Division and the mechanized 4th Infantry Division is challenging the authors' argument in serious, though less obvious, ways. While the authors come up with a lot of anecdotes and selected stories about the two units to support their argument, they fail to assess data critical to the outcome of any combat mission – the casualty rates. This is especially important in the context of the combat attrition model, proposed by this analysis. The better-connected to the local population 101st, which supposedly had better intelligence and situational awareness due to their light patrolling suffered some 5 attacks per day, on average, losing 84 personnel during their tour of duty. The mechanized 4th, which patrolled with their armored vehicles, encountered reportedly 25 attacks per day on average, suffering 34 casualties during its tour (Lyall and Wilson 2009, 101). Let us allow this data to sink in for a moment – so the light division lost 84 people by encountering 5 attacks a day, on average; while the mechanized division lost 34 people being confronted with 25 attacks a day on average. A very illustrative way to quantitatively compare the efficiency of the insurgent attacks against the two units as well as the losses of the two units is through the damage rate of 84:7. This means that the "inefficient" in

getting intelligence 4th Infantry Division was 12 time more effective in preserving the lives of its personnel, and in repelling insurgents' attacks than the 101st division with a "higher situational awareness". It has to be stressed that the intelligence collected by the two units was of tactical nature mostly and fed into their own operations. We also need to consider that by drawing 25 attacks per day, on average, against itself, the 4th division was also able to increase tremendously the exposure of the insurgent, while maintaining a low exposure itself, thus likely inflicting much more casualties against the insurgent. In contrast, 101st was less successful in increasing the exposure of the insurgents, while suffering a higher casualty toll due to its own increased exposure. It likely inflicted less attrition on insurgents as well. If we use the combat attrition model to analyze the combat effectiveness of the sides involved in this scenario, we actually can conclude that 4th Infantry Division was more effective. Intelligence collection is not done for the sake of collecting it – it has to be instrumentalized with one single purpose in mind – to inflict as much damage to the insurgent, while diminishing its own losses. From that perspective, this particular case greatly undermines the Lyall and Wilson argument and theory. This case rings a small alarm bell in regard to how academics frame their research on salient political issues, sometimes creating the impression of a normative bias.

Moreover, there is an increasing number of signals that the winning of "hearts and minds" in counterinsurgency may be a concept with little empirical support. In a more recent example Cohen (2014b) indicated that "despite the Surge's success, in absolute terms, the United States never won 'hearts and minds' in Iraq." He claims

that even in February 2009 polls revealed that only some 30 percent of local respondents believed US and its allies were doing a “good” or a “very good” job. Less than that – some 26 percent – had confidence in the US forces. To further support his claims, Cohen pointed to a March 2008 poll, which indicated that some 42 percent of the respondents believed the attacks against US forces were justifiable. He makes a convincing case that “hearts and minds” approach has not lent strong evidence for its impact on counterinsurgency success.

Gentry (2010) reinforces these ideas through an assessment of eight counterinsurgency campaigns conducted by Britain in Malaya, Kenya, Cyprus, and Norther Ireland, France in Algeria, Portugal in three of its African colonies, and the United States in Philippines and Vietnam. He, for instance, indicates that in Kenya the British military recruited the surrendered Mau Mau fighters, creating “pseudo-gangs” to operate similar to Mau Mau fighting groups and used them to hunt their former colleagues. This is an obvious enemy-centric strategy. Similar examples can be identified in other insurgency cases. For instance, conceptually, Gentry’s example reflects similarities to the American Phoenix program, where the counterinsurgents operated like insurgents themselves, conducting their own intelligence collection through covert observation and development of local sources.[79] In fact, in line with the combat dynamics model this makes a lot of sense. The counterinsurgent, being engaged in a force attrition race, would behave

[79] Claiming this was a population-centric approach is misleading, because the American troops operated behind the “enemy line”.

optimally by employing such strategies as they considerably diminish its own combat exposure. Even though this is a strategy closer to the subset of selective violence, it is not population-centric but enemy-centric. Its goal is to increase the exposure and the attrition of the opponent and not through the engagement of the population. Some of these cases used local population but selected the groups that were already inimical to the insurgent.

What these examples reveal is another way of operationalizing exposure in favor of the counterinsurgent. As the practice reveals, the counterinsurgent frequently maintains an extremely high level of exposure of its troops as its military bases are known while its patrols and reconstruction teams are operating in the open. A more optimal strategy for him would be to have at least part of its troops operating at a lower degree of exposure, involving insurgent-like tactics against the insurgents. Historical examples from a number of successful insurgencies – 19th century Indian wars, the 1899 Philippine insurgency, Nicaragua, Vietnam - indicate that this approach tends to bring successful results (Cassidy 2004). From the conceptual perspective of the proposed model, guided by exposure imperative, this is not surprising. The creation by the counterinsurgent of hunting groups operating like insurgents can tremendously affect both the insurgent's exposure and its attrition dynamics. It also affects its morale and therefore its effectiveness as well.[80]

[80] This follows a similar logic to the one described in the conventional war chapter. A higher morale of the counterinsurgent allows it to longer stay engaged in the operation against the insurgent. The difference is that due to less intensive combat, the losses of the counterinsurgent are much smaller, thus not affecting considerably its overall level of capabilities. The troops morale though are not an accurate representation of the domestic audience morale.

The insurgent finds it more difficult to target these opposing groups as their exposure is minimal. As a consequence of such actions the insurgent's own exposure increases because its location is being identified more frequently by the counterinsurgent through covert search-and-destroy operations. It is one of the few ways for the counterinsurgent to offset the insurgent's advantage in exposure.

To some extent this approach is being applied by various countries through the use of special forces operations. The British used its SAS elements to conduct covert lengthy surveillance against IRA militants and even conducted discriminate and indiscriminate killings of its members (Gill et al. 2016). The Russian military explored this to decrease its exposure in fighting the Chechen rebels by employing raids, ambushes and airborne insertions to launch surprise attacks (Kuznetzov 2000). However, these actions looked more like a heuristic response, emerging more as an accidental policy than a well understood, designed and implemented strategy. In fact, it is largely an underdeveloped method of reducing the exposure of the counterinsurgent and is rarely considered in a conscious and strategic way. It has significant potential though as it would diminish the ability of the insurgent to inflict damage on the counterinsurgent and "a diminished ability to hurt the enemy, not simply military victory, is a major reason to stop fighting" (Slantchev, 2003).

CHAPTER V

Public Policy Implications

From the very start the thesis stated its objective to pursue a public policy focus. Operationalizing a social science theory into policy is a challenging endeavor and requires a significant effort of its own. It would be overly ambitious and unrealistic to pretend this can be done in the limits of this chapter. What is possible to do though, while also suggesting a minimally sizable policy utility, is to examine a few salient policy topics and evaluate the implications of the theory on the issues they raise.

This chapter will elaborate on a few logically interconnected questions, which have been emerging in the policy debates every now and then. To make sure this is indeed the case I scrutinized the research publications in the area of foreign affairs and security policy of a few prominent think tanks, for the last couple of years. For the United States these are RAND Corporation, Council on Foreign Relations, and Brookings Institution. Selecting the topics that emerged and regularly resurface on the agenda of European policymakers I also looked at the publication history of the

European Council on Foreign Relations, the European Union's Institute for Security Studies, Chatham House, and the International Institute for Strategic Studies.[81]

What does the proposed theory tell us about the practical conduct of modern war? What is its possible effects on the future of warfare? How does the presented theoretical framework allow us to better restrain the use of war as a tool of foreign policy? What are the repercussions of the theory's logic on peacekeeping, if any? And finally, what does the theory suggest us in terms of improving arms control? All these questions are connected logically and conceptually. Understanding military effectiveness or how capabilities can map onto military victory allows us to advance our understanding of these related questions as well.

Modern warfare (including insurgencies)

The developed theory suggests that military effectiveness depends on establishing an asymmetric advantage in exposure. This means reducing to the maximum the exposure of own troops and instantly increasing the exposure of the opponent's military forces, while lethally engaging him. How would this materialize from the conceptual domain into the real world? The most obvious way to do this today is for a combatant to have weapon systems that can hit the opponent before entering the lethal (firing) range of the opponent. This is the simple idea behind it –

[81] To explain the focus on the United States and Europe and avoid accusations of being overly Western-centric, I would like to suggest that US and UE continue to significantly shape and influence the global foreign and security policy agenda.

hit the other side, while it cannot hit you back. Or, inflict significantly more damage than the other side can inflict onto your forces. This seems straightforward and does not strike us as rocket science. However, it is a heuristically-designed solution, without understanding the underlying concept behind it: exposure is fundamental to winning battles and wars. Firepower and accuracy are secondary and enter the battlefield only to the extent exposure opens the imaginary “door” for them. Exposure can reduce the effect of firepower and accuracy to zero or can allow them to manifest themselves to their maximum ability. A few examples were offered in the previous chapters. Some others include the Chinese forces entering North Korea to fight the United Nations troops led by the United States. They marched into Korea and approached the UN troops very stealthily, with almost zero degree of exposure Because of this, while superior in firepower the UN troops were initially unable to use their significant advantage. Exposure is the easiest to maintain by infiltrating individual fighters or small groups from different directions and then aggregating them next to the target. Then this aggregated force can strike, disperse and leave back to the base in small groups and through different routes. This logic was to some extent explored by the Russian military when taking over the Ukrainian Crimea and when attacking the Ukraine’s Donbas region. While not necessarily understanding the underlying logic that I elaborated and showed in this thesis, humans have occasionally used it. It is also an evolutionary adaptation strategy among some animals, such as schools of fish avoiding predators (Pitcher and Wyche 1983, Magurran and Pitcher 1987, Magurran 1990).

The effort to break contact with the enemy, while maintaining the ability to target that opponent is one optimal strategy to increase the probability of winning. There are a few other intuitions that the model suggests. This is related to the speed of attrition, and the ability to increase it. In fact, the model and its illustration in the insurgency chapter suggest that the key to victory in these types of combat interaction is the speed of attrition. The quicker one side annihilates its opponent, the lower her own losses are and the more vulnerable and operationally incapable the opponent becomes. To achieve maximal speed of attrition, due to constraints of our physical world, one must employ maximum vigor and aggressiveness resulting in the exercise of maximum violence. It refers to any subset of warfare, be it conventional war or insurgency war.

This is much more counterintuitive, in particular, in the way counterinsurgency is conducted today. If the goal is to defeat an insurgency the presented theory strongly suggests this is possible only through high impact and high speed of attrition. It is under these conditions that the insurgent does not have the required time, necessary to absorb and incorporate the assistance from his supporters, compensating the small (in scale) losses he usually suffers. The insurgency tactics of reduced exposure emerged evolutionary, as a competitive behavior, specifically because it allows it to maintain reduced levels of losses and a low attrition speed. What explains this, to a large extent, is the high ability of the insurgent to control the frequency of its combat interactions, as a rule. Therefore, the counterinsurgent should engage the insurgent more actively, should seek it, to

keep the attrition constant and increasing (see the dynamic on fig. 4.4) despite the fact that the insurgent will try to avoid frequent contacts. It is because of these traits of the insurgency combat dynamics that actions like U.S. Phoenix project in Vietnam, the Mau Mau defectors organized by the British in Kenya,[82] or the Russian special forces and assault teams stealthily seeking the Chechen rebels through their area of operation, were so effective. One thing that can improve in the modern counterinsurgency, in a short time-span, is adjusting both the strategy and its technical support by moving away from a high-exposure one based on patrols and "control", towards a low-exposure one based on stealthy search-and-destroy teams. Patrolling does little to defeat the insurgency, but provides the insurgent with more targets to attack, which consequently would encourage more support from the insurgency's sponsors. Patrols can be used as bait for ambushes, but insurgents will quickly adapt. Patrolling is better to be conducted with unmanned aircrafts and vehicles, while human soldiers should hunt. The counterinsurgent should fight insurgents by employing insurgent own tactics, thus reducing its troops exposure. At equal levels of exposures, the combatant with more resources and firepower will win. Besides, when the counterinsurgent reduces its exposure, this will force the insurgent to elevate its own exposure, because seeking targets to attack and show activity is one of the insurgent's reasons of existence. This change of tactics and at

[82] While both in the Vietnam and Kenya conflicts the counterinsurgents, consequently, took the political decision to withdraw, these specific actions allowed them to significantly increase the insurgents' attrition rates and created conditions for counterinsurgents' military superiority or victory.

the end, strategy, will require political courage and military operational brilliance to implement it effectively.

Another way to view insurgencies (and any type of modern combat) today is that they operate in comparatively resource-scarce environments. Warfare aggravates this, because it destroys resources but also is itself a resource-intensive social activity. This condition threatens an insurgency's "reproduction" and consequently its survivability. In the research and policy discourse on insurgencies we mystified them, empowering phenomena like population support with exaggerated effectiveness. Instead, a following claim merits consideration: it is not the insurgents who exploit the population to advance their goals. It is the population that instrumentalizes the insurgency to remove a governing body that it does not like. The existence of such a preference among a large (or critical) part of the population determines the power, effectiveness and survivability of the insurgency. An existence test would be to conduct free and fair elections and examine whether the population would vote out of power the incumbent against which an insurgency mobilized.

When this (population-based) enabling condition does not exist, the insurgency either does not emerge at all, or perishes quickly. It is most likely a selection effect that we see only insurgencies that had the conditions favorable for their success. The reason for this selection is that insurgencies, being resource-intensive activities, demand considerable and continuous support. In a state of the world where insurgency is strong, this is a signal that the population supports part

of its agenda. It is therefore futile to try to dissuade the population from supporting the insurgency. Unless the population is offered the opportunity to vote in free and fair elections, the only solution is to starve the insurgency through high-speed attrition generated by intense combat interactions, which are coupled with blocking of supplies to the insurgent.

Instead of a statistical perspective on insurgencies we shall allow for a deterministic-evolutionary one. An insurgency materialized specifically because the conditions on the ground favored its emergence and consolidation. One way to see it is by assuming that any social context has individuals unhappy with the system, but not every social context would encourage and maintain successful rebellions. Conditions for the success of a rebellion need to appear. In response, a counterinsurgent has to change that environment, which is facilitating for the insurgent, by making it unfriendly for the evolution of insurgency. One needs to view the insurgency like a living organism and evaluate what makes it survive in that particular social environment. An important condition is access to resources, which was considered and clarified earlier in the chapter. By either being denied the flow of resources or being unable to compensate the quickly deteriorating capabilities with this flow, the insurgency will vanish by itself as its favorable environment is aggressively being changed by the counterinsurgent.

I spend so much space on insurgencies because this is likely to define the way warfare modern warfare will evolve. This will eventually be possible by exploring the potential of exposure and of the attrition speed.

The future of warfare and swarming

What is its possible effect of the proposed theory on the future of warfare? The implementation of the exposure-attrition causal logic with a focus on the speed of attrition would demand advances in technology. One of the reasons the evolution of human warfare focused on firepower, accidentally or not, is because this is the easiest solution. Covering an area with projectiles is an attempt to increase the exposure of an opponent that cannot be targeted more accurately. It is the preference for accidental rather than targeted exposure.

With advances in technology and in particular the developments in the field of Artificial Intelligence, it will be possible to algorithmize the implementation of reduced exposure for one's forces, while increasing that of the opponent. This will lead to the application of the technologies used in missiles to seek-and-destroy airplanes, in unmanned combat platforms, covering all three combat environments on earth. The implementation of this strategy will require a high-degree of automation since the human reaction is unable to optimize the combat constant competition for reducing own exposure while maximizing the one for the opponent. This advance will repeat the history of information processing algorithms, as one percent advantage in reducing exposure of an AI-guided system will offer exponential advantage in attrition capabilities.

A revolution in warfare, in this sense, will not emerge though soon – the operationalization of this combat tactics will require technological developments in sensors, mechanics capable of quickly changing vectors of movement for the unmanned vehicles, technologies of signature reduction, small yield strikes with high precision and limited lateral damage, robotics with tactical autonomy, and engineering solutions to combine all these elements in a complex combat system.

The role of human operator will be critical at the operational and strategic level, assigning the tasks to the AI, which will implement it tactically. Combat, besides attrition, will continue to be about deception and the AI-lead combat platforms will not be soon able to build strategies and operationally guide their implementation in a highly adaptive way. Humans will be indispensable for conducting this task, encouraging the education and training of officers and military thinkers that can operate at a conceptually different level.

A very important element that will be addressed in a following sub-section is the development of an ethical filter for AI-guided combat systems, aiming to restrict specific activities, such as targeting civilians, civilian infrastructure, or other elements of the battlefield that we would like to be protected. In fact, the automatization of warfare will increase our ability to be discriminating in the exercise of violence if the human operator will desire so.

One could recognize in the description above the logic of swarming, which combat application was researched by RAND Corporation, among others (Arquilla and Ronfeldt 2000, Edwards 2005). While not very intuitive this thesis suggests that

swarming can and will be applied in all types of combat interactions, types of warfare, and physical environments. This is what Arquilla and Ronfeldt (2000) suggested as well, claiming that swarming will eventually apply across the entire spectrum of conflict. This includes conflicts from low to high intensity as well as conflicts ranging from civic-oriented missions to military combat operations on land, in the air, and at sea.

To realize that the concept is still in its infancy we just have to look at the definitions different authors apply to swarming. Edwards (2005, 21) argues that swarming "is more like an operational concept or tactic… is just one behavior among several (including guerrilla tactics) that can be exhibited by non-linear, dispersed forces". Arquilla and Ronfeldt (2000) view swarming as a "deliberately structured, coordinated, strategic way to strike from all directions, by means of a sustainable pulsing of force and/or fire, close-in as well as from stand-off-positions" emphasizing dispersion. The logic of swarming, as suggested by this research, is closer to Arquilla and Ronfeldt because it views it as a strategic adaptation and change of warfare. It requires a total change of thinking and a paradigm adaptation from firepower and accuracy towards reducing lethal exposure and increasing the opponent's speed of attrition.

To see other differences, we could point out that Edwards understanding of swarming indicates that the key attributes of a swarm are elusiveness, standoff capability, and superior situational awareness. He also demands that swarms should strive for two types of behavior: encirclement and simultaneity (2005, 113).

The closest variable he has to exposure is what he calls “elusiveness”, which in his model is a higher order variable that is enabled by traits like speed and concealment (Edwards 2005, 113-114). Exposure would be reduced by elusiveness – so we again observe that the focus of researchers is on tools and actions, rather than on principles and dynamics. Also, it is important to emphasize that for both RAND studies the accent is on information, which is again a tool. My call is to focus on exposure and speed of attrition, which are the process and the goal of any effective combat interaction.

In public policy terms, I view the application of swarming and the change of warfare differently – as I view AI as an intermediary between the human operator and the unmanned combat units. Effective swarming with manned combat platforms is not possible, unless we reach a technical breakthrough and connect the human operators’ brains to a technical interface communicating with AI. There is a myriad of issues with that, among which ethical ones are not the most difficult. Even though that sounds somewhat fantastic these are some of the main solutions, which would enable us to maximally explore the concept of evasion of opponent’s lethal action.

However, at more primitive and significantly less efficient levels, the principle is being already used albeit not necessarily conceptualized the way I did. For instance, the Russian General Staff of the Ministry of Defense reported that its opponents in Syria have been using tactics resembling the behavior of a bee swarm. The Russian officials pointed out that this included multiple, dispersed in space and time attacks by small groups of rebels against the Syrian governmental forces,

which nevertheless seemed well coordinated, and were exploring the effect of surprise. As a result of this tactics, the rebels achieved the effect of the permanent, continuous attack, which "exhausted, demoralized, and deprived of initiative even large military formations".[83] It is beyond the scope of this thesis to explore more in-depth the topic of swarming. However, from a policy perspective the theoretical framework about combat dynamics that is developed in this thesis is ideally implemented only through a technological platform like swarming. This is because the optimization of the effect of reduced exposure requires mathematic precision, technical speed, and unhuman coordination. Even with the existing human limitations it is still possible to get significant advantage on the battlefield by implementing the findings of this research, precisely because this is not how military practitioners, researchers and policymakers have thought about the issue.[84]

Arms control and International Humanitarian Law

By developing a theory that examines war outcomes and thus is evaluating military effectiveness, this work contributes to the body of theoretical literature that

[83] Tass.ru. "Genshtab: Boeviki v Sirii ispol'zujut taktiku 'pchelinnogo roja'" [Russian General Staff: Rebels in Syria explore the 'bee swarm' tactics], August 25, 2017, https://tass.ru/armiya-i-opk/4507999.

[84] Swarming may seem, in perspective, as a cheap way of waging war, given machines would replace soldiers. This intuition may not necessarily be accurate – first of all, there are many countries for whom soldiers are preferred to machines, as they have abundance of labor factor of production and scarcity of capital and knowledge. Also, the production of agile assault drones is going to be a costly endeavor for foreseeable future and replacing them will not be that cheap. As an illustration, consider the 2011 NATO bombing of Libya when several NATO member-countries that contributed to the operation run out of munitions and had to purchase more to replenish their stocks.

advances our understanding of a very important phenomenon of human activity – organized coercive violence. In my thesis' dedication section, I wrote that knowledge, similar to a sword, is double-edged. However, similar to how we are unable to cure a disease without understanding how it attacks the human body, we are unable to diminish the frequency of wars without understanding what makes the warfare an effective tool of country's foreign policy. It was the major guiding force for this research to understand better the nature of war in order to elaborate policies that could make it less attractive as a policy tool.

This is not a new intuition – in fact most prominent anti-war efforts and policies came in response to a better understanding of warfare. The body of International Humanitarian Law (IHL) contained to a large extent in four Geneva Conventions of 1949 and two Additional Protocols of 1977 has as its main goal "to limit the effects of armed conflict… for humanitarian reasons".[85] IHL seeks to restrict the use of chemical and biological weapons through its 1972 Biological Weapons Convention and its 1993 Chemical Weapons Convention, among other legal documents. It prohibits the use of weapons that are by nature indiscriminate, unable to differentiate between military and civilian targets. While these rules have a strongly pronounced humanitarian focus, in practice their effect is to limit the effectiveness of some very destructive weapons. For instance, let us examine the case of the use of cluster munitions in inhabited areas, which would be

[85] International Committee of the Red Cross. "What is International Humanitarian Law?" Legal Factsheet, December 31, 2014, https://www.icrc.org/en/doc/assets/files/other/what_is_ihl.pdf. Behind the listed legal documents, the IHL also includes a number of other international agreements, such as 1997 Ottawa Convention on anti-personnel mines.

indiscriminate according to IHL. Its military effect is to tremendously increase the exposure of a combatant that has dispersed throughout a battlespace and explores cover and concealment. Similarly-effective weapons, from the military point of view, designed to target such well-concealed combatants are the incendiary ones, which are also restricted by the IHL. These restrictions are expected to lead to a decrease in the frequency of their use. As a result, the attrition process of the combatants becomes slower, the war can drag on for longer and the costs of supporting the war become higher for belligerents. Observing such political constraints to warfare, policymakers are more likely to reach the conclusion that war is an excessively expensive tool of foreign policy. In fact, while contemplating to write this thesis one of my working hypotheses was that IHL had a decreasing impact on the frequency of interstate wars, due to prohibition of certain weapons. I later on became aware of studies that offered support to this intuition. For instance, empirical observation suggests that classical infringements against another country's sovereignty, such as the military territorial conquests, became less frequent after 1945 (Lacina et al. 2006), partially due to the costs imposed by the body of international norms regulating interstate warfare such as UN Charter (Fazal 2012). This view goes well along with the claim that armed territorial conquests became prohibitively costly (Fearon 1995) political enterprises, under certain conditions.

These allow us to link the better understanding of war dynamics or of conditions that make military conquest more effective to the idea that by restricting this effectiveness could make war costlier. IHL should be complemented with legal

provisions that prohibit or restrict combat tactics and weapons which have a high capacity to diminish the exposure of combatants.[86] Under this condition, the logical chain described earlier would have a higher chance to materialize. In fact, it would make sense to restrict any action that makes war an effective tool of foreign policy. In terms of arms control, it is necessary to restrict weapons that through their wide coverage and indiscriminate targeting are able to reduce exposure of combatants dispersed through a battlespace. It is when such weapons available that a belligerent can apply the principle of indiscriminate violence as a substitute for his insufficient intelligence about the enemy. Another possible use, as suggested by the theory, is that states would use indiscriminate violence to reduce their own exposure and thus attrition rate. Weapons should not be examined based on their purported offensive or defensive qualities but through their ability to reduce the opposing forces exposure.

Peacekeeping

A good working definition of peacekeeping is "the prevention, containment, moderation and termination of hostilities between or within states, through the medium of a peaceful third-party intervention, organized and directed internationally, using multinational forces off soldiers, police and civilians, to restore

[86] Such efforts could trigger the problem of enforcement. As described earlier in the text, the enforcement could be possible if the more powerful countries would be interested to stop the proliferation of these exposure-reducing military technologies that could make weaker actors more able to challenge their interests abroad.

and maintain peace".[87] UN peacekeeping evolved throughout years: from very soft mandates with reduced opportunity to use force towards more robust rules of armed engagements. UN peacekeepers also became a target of continuous and increased attacks.[88]

There has been mounting pressure from the UN military experts on the political side to allow for more robust mandates and rules of engagement that are not restrictive in the use of lethal force. While being more resistant to these calls in the past, in a recent evolution in perceptions about its peacekeeping missions the UN member states and the organization's senior leadership have worked slowly to fix the issue. For instance, they adopted the Resolution 2098 (2013), though which the UN Security Council enabled the establishment of a first-ever "offensive" combat force to "neutralize and disarm" Congolese rebels and foreign armed groups.[89] A 2017 report commissioned by UN Secretary General Guterres, with Brazilian Lt.Gen. Carlos Alberto dos Santos Cruz as the lead author, also argued for a more robust and aggressive stance of the peacekeeping missions. It makes the telling suggestion that "nobody attacks a stronger opponent" and explains that there is a major problem in the way UN missions are planned. How are the findings of this

[87] Quoted in Last (1995).
[88] According to a report commissioned by UN Secretary General Antonio Guterres, since 1948 over 3,500 U.N. personnel lost lives in organization's peace operations. During the 2013-2017 period, 195 UN personnel were killed. See Carlos Alberto dos Santos Cruz, William R. Phillips, Salvator Cusimano. "Improving Security of United Nations Peacekeepers: We Need to Change the Way We Are Doing Business." December 19, 2017. https://peacekeeping.un.org/sites/default/files/improving_security_of_united_nations_peacekeepers_report.pdf
[89] See UN Security Council 6943rd Meeting Press Release. "'Intervention Brigade' Authorized as Security Council Grants Mandate Renewal for United Nations Mission in Democratic Republic of Congo." March 23, 2013, https://www.un.org/press/en/2013/sc10964.doc.htm.

thesis useful in understanding this major challenge and in proposing both feasible and effective solutions?

First of all, terms of the nature of the problem, the theory developed in this work supports conceptually the conclusions of the Cruz report. If an international armed mission that is mandated to reduce and prevent violence, and support a peace process would like to be effective, it has to be able to understand the dynamics of violence in its area of operations. The mission's mandate should be predicated on this understanding. The thesis suggests that a conflict in which two belligerents are engaged in small skirmishes is going to last for a long period, if the rebel group is able to secure access to resources.[90] This is due to the fact that the rebels' attrition is too slow, unable to undermine the re-supplying of lost capabilities by the rebels, in order for the government to win militarily. The peacekeeping mandate should account for this and, if feasible, should consider addressing the rebel source of supplies.

In exercising its duties the mission is very likely to step on some of the armed actors' toes, triggering violent attacks against its personnel. To discourage this, it has to make itself a harder target. The force protection effort should consider increasing the exposure of a potential attacker as he would approach the base. It should plan and organize its patrols with the same idea to elevate the exposure of potential attackers, by using tactical drones, for instance, as well as specific

[90] It is less important whether the resources are acquired through illegal extraction of natural resources, smuggling, or through foreign support.

reconnaissance procedures. The missions should also avoid making themselves soft targets, which is facilitated by their own increased exposure and by allowing for conditions that would produce a high speed of attrition among the mission's troops. While usual rules of engagement aim to decrease the instances when peacekeepers are to use their weapons, this thesis theory strongly suggest a different action. `

The mandate should allow for an extremely aggressive response that would maximize the armed violence locally and in a limited span of time. The burst of violent response by peacekeepers to an attack against them would lead to an increased speed of attrition of the attacker's forces. If the attacker is rational, it would think twice before attacking again – a continuous policy of robust and aggressive local responses would create high attrition rate of the attackers. It will also form a reputation of a peacekeeping force that should not be messed with. This is an optimal response even if the attacker is less rational. Under this condition, continuous attacks against a UN force prepared to respond robustly in terms of tactical combat engagements would lead to attacker's annihilation.

The thesis offers a theoretical context, which can be used to explore and explain both the conduct of violence but also how to avoid and prevent violence. It has a wide application in situations that may require the use of force or confronting an antagonist force. It offers insights on how to make force application efficient and also how to restrain it. In particular, while the UN military experts learned through decades of experience that weak rules of engagement make the peacekeepers vulnerable, this thesis derived accurately the underlying dynamics describing the

violence against UN peacekeepers, the logic of how to curb it, and suggested the strategic incentive of combatants to target UN forces. It offered the justification for a more robust mandate and rules of engagement.

CHAPTER VI

Conclusions

In this intellectual exploration I start my journey from the idea that in order for a policymaker to embark on the risky decision of escalating a diplomatic dispute into war, she would have to make certain assumptions about how military force translates into victory on the battlefield. That is, she would have to understand why in some cases power preponderance can lead to victory, while in others it is the path to defeat. The underlying assumption for this logic is that no policymaker would like to escalate a crisis into war, if she does not have at least somewhat high, even though subjective expectation that her side is going to win. Historically, power preponderance has been the main factor affecting the expectation of victory.

Since there is a considerable subset of conflicts where superiority in military capabilities did not determine a victorious outcome of war, the riddle we want to solve is how exactly military capabilities could lead to victory. A starting point in addressing that question was the assumption that the fundamental war dynamics was guided by the process of attrition of war resources. In order to further understand and clarify this intuition, the thesis identified, formalized and showed that it is exposure - a physical property of combat interaction - and not the firepower

or its accuracy that guides the speed of attrition specifically and attrition generally. This thesis found that exposure determines victory on the battlefield by affecting the speed of attrition. This is the essential mechanism of war summarized in a few sentences.

This research is contributing to the conflict escalation framework by focusing on the spatial constraints of the combat ecology that shape the policymaker's behavioral and motivational factors, following their exposure to that ecology. This approach connects the political science research on crisis escalation, preponderantly focused on behavioral and motivational factors, with the military operations research that is focused mostly on the dynamics of combat environment.

In this research I develop a novel theory about the dynamics of violent conflict and how military capabilities map onto victory. I argue that the key factor driving the probability of victory is not the capabilities-firepower factor or its accuracy, but the degree of lethal exposure of one's forces, which essentially affects these forces' attrition speed. I build a simple model of dynamic combat attrition from which I derive the intuition that capabilities play the role of nothing more than a reservoir, from which a side can supply its combat efforts. A side with a higher number of forces will just have a larger reservoir. Manipulating the degree of exposure requires superior skills in tactics and strategy, or superiority in technology. The effects of exposure therefore are less visible through human perception. When they emerge, there is a tendency to attribute them to some more obvious driver, such as the power preponderance. This is why capabilities, strategy, or technology have been typically

selected as some of the dominant explanations for war outcome, while in fact they are only proxies.

The reason why I singled out exposure as the key explanatory variable is because the degree of exposure of forces can completely block the impact of capabilities and fire accuracy on troops attrition speed, if we decrease exposure through manipulation. It is only when the value of exposure is very high, like in hand-to-hand combat that the role of capabilities and fire accuracy is maximized. Therefore, it is the degree of exposure of one's military forces that affects war outcome. This explains how even a belligerent with considerably lower capabilities can win a combat interaction, if it manages to get an advantage in the level of lethal exposure, comparing to the opponent. I explore in detail the first day of the Somme Offensive to identify these traits in the combat interactions between the allied British and French forces that are opposing the Germans. The intra-case variations in troops' exposure and the resulting differences in speed and degree of attrition is providing a vivid illustration of my theory.

While examining the exposure variable in the context of insurgency warfare, I theorize and generate more support for the effectiveness of the enemy-centric strategy, in contrast to the population-centric strategy of counterinsurgency. In fact, the effect of exposure is even more pronounced in the insurgency context because of the asymmetry in its endowments for the two belligerents. By exploring the differences in Russia's two Chechen campaigns, in particular the varying degrees of exposure of both the Russian forces and the Chechen rebels it is again possible

to illustrate the critical and primary role the exposure plays, in contrast to capabilities and fire accuracy.

Besides its obvious policy relevance, I argue that answering the thesis question makes a few contributions to political science literature that explores crisis escalation. One main contribution is that it bridges two bodies of literature on conflict that speak past each other: the bargaining diplomacy scholarship in political science and the military effectiveness literature from military operations research. Another contribution of this analysis is the introduction of a different optics on military effectiveness, suggesting a dramatic shift from preponderance of force and firing accuracy towards the ability to avoid targeting, through the variation of troops exposure to lethal fire. This advances a rather different perspective, as leading relevant political science (and military operations) research has been built on the power preponderance assumption.

In light of these findings we can limit the role of capabilities to that of a reservoir, while elevating the role of troops' degree of exposure. If the exposure rate of combatants does not differ much, then obviously the side with a larger reservoir of war resources is likely to remain with some leftover troops after the battle. However, if the difference in exposure exists and the side with the bigger reservoir loses troops at higher speed, the preponderance of force advantage can very likely be undermined. Lower exposure and a resulting lower speed of attrition of forces are the factors that affect war outcome more strongly than capabilities and their fire effectiveness.

The suggestion that exposure can affect the probability of victory and the speed of attrition of war capabilities also implies that exposure can affect *the costs* of violent conflict. The longer two sides are fighting, the more resources they are destroying, both combat-related and those affecting their economic development.

Another important insight is that this analytic exercise undermines the role of capabilities in determining the probability of military victory. I show that both the superior number of forces and their higher firing effectiveness may be to a very large extent dependent upon the exposure of the opponent's troops, which I define and introduce. This suggests that fire effectiveness, which has been one of the two dominant variables in the military operations research to explain probability of victory, is in fact a proxy and is very much dependent on the degree of opponent's troops exposure. To put it differently, when troops are in direct contact, given the mass-effect of modern munitions and the spatial nature of combat, fire accuracy is secondary and fully controlled by exposure.

The research suggested the less intuitive idea that even in conventional war, including some of its most intense forms like trench warfare, the variable of exposure can play a critical role and make the difference between victory and defeat. It pointed out that the classical candidates affecting defense and security policy, but also guiding academic debate – capabilities, fire accuracy, and morale – can make an impact on the level of attrition of troops only through exposure. In fact, their effects can be blocked to the minimum by manipulating the level of exposure. Examining both older and more recent case studies revealed that military planners and

commanders did not even conceive battlefield operations in terms of exposure. For instance, in the Somme Offensive campaign the British military commanders erroneously assumed that superiority in artillery and infantry were sufficient conditions to defeat the opposing German forces. While they obviously viewed attrition as a dependent variable in their strategizing, the capabilities and the resulting firepower were their key and only independent variables of interest. The Russian generals, by marching their troops in parade formations into Grozny when starting their first Chechen campaign, have also bet on superior capabilities. In the second Chechen campaign they did not change the perception about the role of capabilities, though they intuitively started to manipulate the size of the Russian troops' exposure through increased firepower, which also increased the rebels' exposure.

While exploring the exposure variable in the context of insurgency war, I argue that the widely-popular both in the academia and policymaking dichotomy of population-centric and enemy-centric approaches to counterinsurgency is misleading and fails to completely account for the imperatives of the combat attrition dynamics. An important stepping stone for that understanding was the acceptance of the physical fact that any combat, at the point of contact, resumes to mutual attrition. I argued and showed that insurgency is conceptually a subset of conventional warfare and therefore is influenced by the effects of combat exposure and attrition. By evaluating the Russian counterinsurgency efforts during the second

Chechen campaign, I indicated that the operationalization of exposure in various policy decisions is possible, and it can take multiple forms and shapes.

The the Russian preference to put emphasis on wining “hearts and minds” of its own population (as opposed to the population linked to the insurgent), thus aiming to reduce the costs of its brutal tactics used in Chechnya, among other things, revealed both a richness of relevant political solutions and the limitations of the existing debate. The Russian application of brute force strategy in counterinsurgency, as it emerged from the selected case, revealed that under certain conditions this is both a rational and an optimal strategy, considering the payoff function of the Russia’s leadership. Since the highest priority of the Russian military was to reduce its combat losses and avoid politically-costly humiliation of defeat, it did not even train its troops to engage in urban warfare. It instead suppressed the rebels in urban terrain from distance and through an indiscriminate application of superior force.

Similar tactics have been recently observed in Syria, being employed by both Russia’s and Assad’s forces. They were explored in some of the Iraqi government’s operations against ISIS, in particular in the 2016-2017 battle of Mosul. This is a striking deviation from the previously practiced and self-imposed restraint on the use of massive firepower in inhabited areas by the US and NATO troops in Afghanistan. It shows that in a coalitional setting, when the Western forces operate in support of

a local government, they can transfer the attribution for the application of enemy-centric, indiscriminate use of force to that local actor.[91]

Furthermore, I show that there exists a set of imperatives, which establishes a dependency of the fighting capabilities of an insurgent and even its survival on two interrelated conditions: i) its ability to evade or diminish the degree and speed of combat attrition, inflicted by the counterinsurgent; and ii) its ability to compensate this combat attrition through extracting resources from the local population and/or foreign sponsors. In order to survive, the insurgent would have to, at least, keep the two quantities equal. Or, if the insurgent wants to maintain any chance for success, it would have to keep the attrition level well below the outside compensating resource flows. Viewing insurgency war from the counterinsurgent's perspective, the secret to success is to sustain an effective speed of attrition of the insurgent forces, through combat or interdiction methods. It has to be of a magnitude and intensity, which the insurgent cannot compensate through supplies and resources that it receives from its local and foreign supporters. It is then that specific balance between these two quantities, which can determine the violence threshold at which the indiscriminate violence (brute force) becomes effective in suppressing the insurgent through attrition solely. This logic significantly elevates the role and importance of combat attrition in insurgency wars, which lately has been neglected in favor of strengthening indigenous population support.

[91] For pertinent details on the battle over Mosul, see BBC, "How the Battle for Mosul Unfolded," 10 July 2017, https://www.bbc.com/news/world-middle-east-37702442.

Another important detail to emphasize is the logic that guides the supply of the resource flows for the insurgent, in contrast to the aggregation of military reserves and supplies in conventional combat. If in conventional combat reserves are a part of the order of battle and they are joining the combat, as the operational needs emerge, the insurgent is by definition facing a scarcity of resources. The support it receives from the indigenous population or foreign sponsors is being integrated into its overall capacities and operations in a much slower fashion. This specific supply dynamics together with insurgency's smaller size, makes the insurgent particularly sensitive to combat attrition, which usually results in considerable losses of its scarce manpower and resources.

By applying the logic of combat exposure, the thesis reframed some contradictory interpretations of counterinsurgents' tactics, viewed as the application of indiscriminate violence. The exposure variable provides the suitable theoretical framework necessary to explain alternative interpretations of indiscriminate violence (as substitute for intelligence). One of them is the effort of the counterinsurgent to reduce its own troops' combat exposure, while attempting to maximize the attrition impact on the insurgent troops. It can achieve this by accidental exposure - saturating a wider area with massive firepower in which the counterinsurgent believes the insurgent to be located, though not knowing its exact location. It is a logic that justifies conceptualizing exposure in spatial terms of being either *targeted or accidental*. Using the terminology and concepts from the first chapter, the counterinsurgent is thus trying to aggregate a few adjacent battlespaces to prevent

the dispersion of its capabilities and the erosion of this advantage, while attempting to maintain a high *speed of attrition* of his opponent. It confirms the conceptual intuition of the theory that this work introduces, which also implies exposure is a spatial and dynamic concept requiring specific methodological treatment.

I argue that the research findings provide important insights for public policy, both in terms of military policy and war-prevention. From a military operational point of view it suggests that it makes less sense to invest into new and more powerful weaponry, as higher combat efficiency would be achieved by developing evasion technics and specific technologies of combat, such as long-range strikes as well as intelligence gathering to quicker and better identify the opponent's location. The model indicates that preponderance of force may have an impact on the probability of victory, but only in close-range stationary combat, and at the price of a higher level of attrition to both belligerents. Due to modern technological advances, accuracy of troops is less important in conventional warfare where non-discriminatory and mass-destruction munitions such as multiple-launch rocket systems and artillery strikes are used. Fire effectiveness continues to be useful in small-scale ambushes and close-quarter combat of infantry, though it is also highly dependent on the conflicting sides' force exposure.

From arms control and peacekeeping perspectives it is important not only to design and implement international legally-binding measures that restricts non-discriminatory munitions or the targeting of inhabited areas, but also further strengthen such existing measures. This would lead to the decreasing exposure of

a defender, which will make it costlier for an offender to wage war. It should, in general, make war a more expensive tool of foreign policy and encourage states to more frequently resort to diplomatic measures of conflict resolution.

The proposed explanation of how military capabilities can transform into victory provides additional opportunities for military strategists and planners, for defense industry, arms control and peacekeeping organizations, as well as policymakers. These opportunities stem from the fact that exposure, as an independent variable in the initial model, can be transformed into a dependent variable. The policymaker then could enhance national defense and security by designing policies that aim at preparing the national military for its combat exposure's minimization, as the need arises.

References

Abbott, Frank J. 1994. *The Greek Civil War, 1947-1949: Lessons for the Operational Artist in Foreign Internal Defense*. Monograph. Kansas, Fort Leavenworth: School of Advanced Military Studies, https://apps.dtic.mil/dtic/tr/fulltext/u2/a284240.pdf.

Achivarina, Vera and Simon F. Reich. 2006. "No Place to Hid: Refugees, Displaced Persons, and the Recruitment of Child Soldiers." *International Security* 31 (1): 127-164.

Adamson, P.B. 1977. "A Comparison of Ancient and Modern Weapons in the Effectiveness of Producing Battle Casualties." *Journal of the Royal Army Medical Corps* 123 (2): 93-103.

Alesina, Alberto and Enrico Spolaore. 2005. "War, Peace, and the Size of Countries." *Journal of Public Economics* 89 (7): 1333-1354.

Ancker, C.J. 1995. "A Proposed Foundation for a Theory of Combat." *Naval Research Logistics* 42 (3): 311-343.

Arena, Philip. 2008. "Success Breeds Success? War Outcomes, Domestic Opposition, and Elections." *Conflict Management and Peace Science* 25 (2): 136-151.

Arquilla, John, and David Ronfeldt. 2000. Swarming and the Future of Conflict. RAND Documented Briefing DB-311-OSD. Santa Monica, CA: RAND Corporation, https://www.rand.org/content/dam/rand/pubs/documented_briefings/2005/RAND_DB311.pdf.

Arreguin-Toft, Ivan. 2001. "How the Weak Win Wars: A Theory of Asymmetric Conflict." *International Security* 26 (1): 93-128.

Balcells, Laia and Stathis N. Kalyvas. 2014. "Does Warfare Matter? Severity, Duration, and Outcomes of Civil Wars." *Journal of Conflict Resolution* 58 (8): 1390-1418.

Barabas, Jason. 2004. "How Deliberation Affects Policy Opinions." *American Political Science Review* 98 (4): 687-701.

Bas, Muhammet A., and Roert J. Schub. 2016. "How Uncertainty about War Outcomes Affects War Onset." *Journal of Conflict Resolution* 60 (6): 1099-1128.

Bates, Robert H., Avner Greif, Margaret Levi, and Jean-Laurent Rosenthal. 1998. *Analytic Narratives.* Princeton, NJ: Princeton University Press.

Bellamy Ronald F. and Russ Zajtchuk (eds.). 1991. *Textbook of Military Medicine. Part 1: Warfare, Weaponry, and the Casualty. Conventional Warfare: Ballistic,*

Blast, and Burn Injuries (Vol.5). Washington DC: U.S. Government Printing Office.
Berman, Eli, Jacob N. Shapiro and Joseph H. Felter. 2011. "Can Hearts and Minds be Bought? The Economics of Counterinsurgency in Iraq." *Journal of Political Economy* 119 (4): 766-819.
Berman, Eli, Joseph H. Felter, and Jacob N. Shapiro. 2018. *Small Wars, Big Data: An Information Revolution in Modern Conflict.* Princeton: Princeton University Press.
Biddle, Stephen D. 2001. "Rebuilding the Foundations of Offense-Defense Theory." *Journal of Politics* 63 (3): 741–774.
Biddle, Stephen D. 2002. *Afghanistan and the Future of Warfare: Implications for Army and Defense Policy*. Carlisle, PA: Strategic Studies Institute, U.S. Army War College.
Biddle, Stephen D. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.
Biddle, Stephen, Jeffrey A. Friedman, and Jacob N. Shapiro. 2012. "Testing the Surge: Why did Violence Decline in Iraq in 2007?" *International Security* 37 (1): 7-40.
Bjørkhaug, Ingunn. 2010. *Child Soldiers in Colombia: The Recruitment of Children into Non-State Violent Armed Groups*. MICROCON Research Working Paper 27, Brighton: MICROCON.
Blainey, Geoffrey. 1973. *The Causes of War*. New York, NY: Free Press.
Boff, Jonathan. 2014. "The Morale Maze: the German Army in Late 1918." *Journal of Strategic Studies* 37 (6-7): 855-878.
Boeke, Sergei and Bart Schuurman. 2015. "Operation Serval: A Strategic Analysis of the French Intervention in Mali, 2013-2014." *Journal of Strategic Studies* 38 (6): 801-825.
Bonder, Seth. 2002. "Army Operations Research – Historical Perspectives and Lessons Learned." *Operations Research* 50 (1): 25-34.
Bonds, Timothy M., Michael J. Mazarr, James Dobbins, Michael J. Lostumbo, Michael Johnson, David A. Shlapak, Jeffrey Martini, Scott Boston, Cristina L. Garafola, John Gordon IV, Sonni Efron, Paul S. Steinberg, Yvonne K. Crane, and Daniel M. Norton. 2019. *America's Strategy-Resource Mismatch: Addressing the Gaps Between U.S. National Strategy and Military Capacity*. Report RR-2691. Santa Monica, CA: RAND Corporation, 2019. https://www.rand.org/pubs/research_reports/RR2691.html.
Boorman, Howard L. and Scott A. Boorman. 1966. "Chinese Communist Insurgent Warfare, 1935-49." *Political Science Quarterly* 81 (2): 171-195.
Britt, Thomas W., James M. Dickinson, DeWayne Moore, Carl A. Castro, Amy B. Adler. 2007. "Correlates and Consequences of Morale Versus Depression Under Stressful Conditions." *Journal of Occupational Health Psychology* 12 (1): 34-47.
Bueno de Mesquita, Bruce. 1981. *The War Trap*. New Haven, CT: Yale University Press.

Bueno de Mesquita, Bruce, and Randolph M. Siverson. 1995. “War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability.” *American Political Science Review* 89 (4): 841-855.

Bueno de Mesquita, Bruce. 2000. *Principles of international politics*. Washington, DC: CQ Press.

Byman, Daniel, Peter Chalk, Bruce Hoffman, William Rosenau, and David Brannan. 2001. *Trends in Outside Support for Insurgent Movements.* Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1405.html.

Cassidy, Robert M. 2004. “Winning the War on the Flea: Lessons from Guerrilla Warfare.” *Military Review* 84 (5): 41-46.

Ciziunas, Pranas. 2008. “Russia and the Baltic States: Is Russian Imperialism Dead?” *Comparative Strategy* 27(3): 287-307.

Clausewitz, Carl von. 1832/1984. *On War. Princeton*, NJ: Princeton University Press.

Cohen, Raphael S. 2014a. “Just How Important Are ‘Hearts and Minds’ Anyway? Counterinsurgency Goes to the Polls.” *Journal of Strategic Studies* 37 (4): 609-636.

Cohen, Raphael S. 2014b. *“Beyond ‘Hearts and Minds’.”* PhD diss., Georgetown University.

Dannreuther, Roland, and Luke March. 2008. “Chechnya: Has Moscow Won?” *Survival* 50 (4): 97-112.

Desch, Michale C. 2019. “How Political Science Became Irrelevant: The Field Turned its Back on Beltway.” *The Chronicle of Higher Education*, February 27.

Deitchman, Seymour J. 1962. “A Lanchester Model of Guerrilla Warfare.” *Operations Research* 10 (6): 818-827.

Deverell, Christopher M. 2005. “X. Haig versus Rawlinson– Manoeuvre versus Attrition: The British Army on the Somme, 1916,” *Defense Studies* 5 (1): 124-137.

Drezner, Daniel W. 2000. “Ideas, Bureaucratic Politics, and the Crafting of Foreign Policy.” *American Journal of Political Science* 44 (4): 733-749.

Druckman, James N. and Arthur Lupia, “Preference Formation,” *Annual Review of Political Science*, Vol.3 (June 2000), pp.1-24.

Duyvesteyn, Isabelle. 2011. “Hearts and Minds, Cultural Awareness and Good Intelligence: The Blueprint for Successful Counterinsurgency?” *Intelligence and National Security* 26 (4): 445-459.

Edmonds, James. 1932. *Military Operations: France and Belgium, 1916*. London: Macmillan and Co Ltd.

Edwards, Sean J.A. 2005. Swarming and the Future of Warfare. RAND Dissertation, RGSD-189, https://www.rand.org/pubs/rgs_dissertations/RGSD189.html.

Epstein, Joshua M. 1985. *The Calculus of Conventional War: Dynamic Analysis without Lanchester Theory*. Washington, DC: The Brookings Institution.

Egnell, Robert. 2010. “Winning ‘Hearts and Minds’? A Critical Analysis of Counter-Insurgency Operations in Afghanistan.” *Civil Wars* 12 (3): 282-303.

Eichenberg, Richard C. 2005. “Victory Has Many Friends: U.S. Public Opinion and the Use of Military Force, 1981-2005.” *International Security* 30 (1): 140-177.
Falkenrath, Richard A. 1995. *The Origins and Consequences of the CFE Treaty*. Cambridge, MA: MIT Press.
Fearon, James D. 1995. “Rationalist Explanation for War.” *International Organization* 49 (3): 379-414.
Fearon, James D. and David D. Laitin. 1996. “Explaining Interethnic Cooperation.” *American Political Science Review* 90 (4): 715-735.
Fearon, James D. and David D. Laitin. 2003. “Ethnicity, Insurgency and Civil War.” *The American Political Science Review* 97 (1): 75-90.
Felbab-Brown, Vanda. 2005. “The Coca Connection: Conflict and Drugs in Colombia and Peru.” *Journal of Conflict Studies* 25 (2): 104-128.
Ferguson, Niall. 2004. “Killing in the Age of Total War: Towards a Political Economy of Military Defeat.” *War in History* 11 (2): 148-192.
Filson, Darren and Suzanne Werner. 2002. “A Bargaining Model of War and Peace: Anticipating the Onset, Duration, and the Outcome of War.” *American Journal of Political Science* 46 (4): 819-837.
Foley, Robert T. 2011. “Learning War’s Lessons: The German Army and the Battle of the Somme 1916.” *The Journal of Military History* 75 (2): 471-504.
Frei, Daniel (ed.). 1982. *Managing International Crises*. Beverly Hills, CA: SAGE Publications.
Galeotti, Mark. 2002. “’Brotherhoods’ and ‘Associates’: Chechen Network of Crime and Resistance.” *Law Intensity Conflict and Law Enforcement* 11 (2-3): 340-352.
Galula, David. 1964. *Counterinsurgency Warfare: Theory and Practice.* Westport, CT: Praeger Security International.
Gartzke, Erik. 1999. “War is in the Error Term.” *International Organization* 53 (3): 567-587.
Gentry, John A. 2010. “Intelligence Learning and Adaptation: Lessons from Counterinsurgency Wars.” *Intelligence and National Security* 25 (1): 50-75.
George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
Gilbert, Bentley B. and Paul P. Bernard. 1959. “The French Army Mutinies of 1917.” *The Historian* 22 (1): 24-41.
Gill, Paul, James A. Piazza and John Horgan. 2016. “Counterterrorism Killings and Provisional IRA Bombings, 1970-1998.” *Terrorism and Political Violence* 28 (3): 472-496.
Glaser, Charles L. and Chaim Kaufmann. 1998. “What is Offense-Defense Balance and Can we Measure It?” *International Security* 22 (4): 44-82.
Griffith, Paddy. 1994. *Battle Tactics of the Western Front: The British Army’s Art of Attack, 1916-18*. New Haven: Yale University Press.
Grinker, Roy R. and John P. Spiegel. 1945. *Men Under Stress*. Philadelphia: Blakiston.
Haas, Mark L. 2001. “Prospect Theory and the Cuban Missile Crisis.” *International Studies Quarterly* 45 (2): 241–270.

Hack, Karl. 1999. “British Intelligence and Counter-Insurgency in the Era of Decolonization: The Example of Malaya.” *Intelligence and National Security* 14 (2): 124-155.
Hashim, Ahmed S. 2013. *When Counterinsurgency Wins: Sri Lanka’s Defeat of Tamil Tigers.* Philadelphia, PN: University of Pennsylvania Press.
Hart, Peter. 2008. *The Somme: The Darkest Hour on the Western Front.* New York: Pegasus Books.
Heather, Randall W. 1990. “Intelligence and Counter-Insurgency in Kenya, 1952-56.” *Intelligence and National Security* 5 (3): 57-83.
Heisenberg, Dorothee. 2005. “The Institution of ‘Consensus’ in European Union: Formal Versus Informal Decision-Making in the Council.” *European Journal of Political Research* 44 (1): 65-90.
Helmbold, Robert L. 1964. “Some Observations on the Use of Lanchester’s Theory for Prediction.” *Operations Research* 12 (5): 778-781.
Hirshleifer, Jack. 2000. “The Macrotechnology of Conflict,” *Journal of Conflict Resolution* 44 (6): 773-792.
Hoffman, Bruce. 2006. “Insurgency and Counterinsurgency in Iraq.” *Studies in Conflict and Terrorism* 29 (2): 103-121.
Hoffman, Bruce. 2007. “Neo-Classical Counterinsurgency?” *Parameters* 38 (2): 71-87.
Hopkins, George E. 1966. “Bombing and the American Conscience During World War II.” *The Historian* 28 (3): 451-473.
Humphreys, Macartan and Jeremy M. Weinstein. 2008. “Who Fights? The Determinants of Participation in Civil War.” *American Journal of Political Science* 52 (2): 436-455.
Iatrides, John O. 1993. “The Doomed Revolution: Communist Insurgency in Postwar Greece.” In *Stopping the Killing. How Civil Wars End*, edited by Roy Licklider, 205-234. New York: New York University Press.
Irincheev, Bair. 2009. *The Mannerheim Line 1920-1939: Finnish Fortifications of the Winter War*. Oxford: Osprey Publishing.
Jervis, Robert. 1978. “Cooperation Under the Security Dilemma.” *World Politics*, 30 (2): 167-214.
Jia, Hao, and Stergios Skaperdas. 2012. “Technologies of Conflict.” In *The Oxford Handbook of the Economics of Peace and Conflict*, edited by Michelle R. Garfinkel and Stergios Skaperdas, 449-472. Oxford: Oxford University Press.
Johns, Robert and Graeme A. M. Davies. 2019. “Civilian Casualties and Public Support for Military Action: Experimental Action.” *Journal of Conflict Resolution* 63 (1): 251-281.
Jones, Seth G, and Patrick B. Johnston. 2012. “The Future of Insurgency.” *Studies in Conflict and Terrorism* 36 (1): 1-26.
Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.

Kalyvas, Stathis N., and Matthew A. Kocher. 2007. “How “Free” is Free-Riding in Civil War? Violence, Insurgency, and the Collective Action Problem.” *World Politics* 59 (2): 177-216

Kalyvas, Stathis N., and Laia Balcells. 2010. “International System and Technologies of Rebellion: How the End of the Cold War Shaped Internal Conflict.” *American Political Science Review* 104 (3): 415-429.

Keegan, John. 1978. *The Face of Battle.* New York: Penguin Books.

Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton, NJ: Princeton University Press.

Kilcullen, David. 2006. “Counter-insurgency Redux.” *Survival* 48 (4): 111-130.

Kilcullen, David. 2009. *The Accidental Guerilla: Fighting Small Wars in the Midst of a Big One*. Oxford: Oxford University Press.

Kraehe, Enno. 1944. “The Motives Behind the Maginot Line.” *Military Affairs* 8: 109-122.

Krause, Jonathan. 2016. “The Evolution of French Tactics 1914-16”. In *The Battle of the Somme*, edited by Strohn Matthias, 199-214. Oxford: Osprey Publishing.

Krause, Lincoln B. 2009. “Playing for the Breaks: Insurgent Mistakes.” *Parameters* 39 (2): 49-64.

Krause, Lincoln B. 1996. “Insurgent Intelligence: The Guerrilla Grapevine.” *International Journal of Intelligence and Counterintelligence* 9 (3): 291-311.

Kress, Moshe. 2012. “Modeling Armed Conflicts.” *Science* 336 (6083): 865-869.

Kuznetzov, Nikolay. 2000. “Novyje Real’nosti Boja.” *Orientir* 6, accessed at www.soldiering.ru/army/file/r_war.php.

Lake, David A. 1992. “Powerful Pacifists: Democratic States and War.” *American Political Science Review* 86 (1): 24-37.

Last, David M. 1995. “Peacekeeping Doctrine and Conflict Resolution Techniques.” *Armed Forces and Society* 22 (2): 197-210.

Lepingwell, John W.R. 1987. “The Laws of Combat? Lanchester Reexamined.” *International Security* 12 (1): 89-134.

Levi, Margaret and Weingast, Barry R. *Analytic Narratives, Case Studies, and Development* (September 6, 2016). Available at SSRN: https://ssrn.com/abstract=2835704 or http://dx.doi.org/10.2139/ssrn.2835704.

Levy, Jack S. 2003. “Applications of Prospect Theory to Political Science.” *Synthese* 135 (2): 215-241.

Lupfer, Timothy T. 1981. *The Dynamics of Doctrine: The Changes in German Tactical Doctrine During the First World War*. Fort Leavenworth, KS.: Combat Studies Institute.

Lupia, Arthur, Mathew McCubbins and Samuel Popkin. 2000. *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*. Cambridge: Cambridge University Press.

Luttwak, Edward N. 2001. *Strategy: The Logic of War and Peace*. Cambridge, MA: Harvard University Press.

Lyall, Jason. 2009. “Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya” *Journal of Conflict Resolution* 53 (3): 331-362.

Lyall, Jason and Isaiah Wilson III. 2009. “Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars.” *International Organization* 63 (1): 67-106.
Lyall, Jason. 2010. “Are Co-ethnics More Effective Counterinsurgents? Evidence from the Second Chechen War.” *American Political Science Review* 104 (1): 1-20.
Lyall, Jason, Graeme Blair and Kosuke Imai. 2013. “Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan.” *American Political Science Review* 107 (4): 679-705.
Lyall, Jason, Yuki Shiraito and Kosuke Imai. 2015. “Coethnic Bias and Wartime Informing.” *The Journal of Politics* 77 (3): 833-848.
Mack, Andrew. 1975. “Why Big Nations Lose Small Wars: The Politics of Asymmetric Conflict,” *World Politics* 27 (2):175-200.
Magurran, Anne E. 1990. "The Adaptive Significance of Schooling as an Anti-Predator Defence in Fish." *Annales Zoologici Fennici* 27 (2): 51-66.
Magurran, A.E, and T.J. Pitcher. 1987. “Provenance, Shoal Size and the Sociobiology of Predator-Evasion Behavior in Minnow Shoals.” *Proceedings of the Royal Society B* 229 (1257): 439-465.
Mahan, Alfred T. 1890. *The Influence of Sea Power Upon History, 1660-1783*. Boston, MA: Little, Brown and Company.
Mackinder, John H. 1904. “The Geographical Pivot of History.” *The Geographical Journal* 23(4): 421-444.
Mackintosh, Malcolm. 1985. “The Russian Attitude to Defense and Disarmament,” *International Affairs (Royal Institute of International Affairs 1944-)* 61 (3): 385-394.
Malkasian, Carter. 2002. *A History of Modern Wars of Attrition*. Westport, CT: Praeger Publishers.
Maoz, Zeev. 1983. “Resolve, Capabilities and the Outcomes of the Interstate Disputes, 1817-1976.” *Journal of Conflict Resolution* 27 (2): 195-229.
Maoz, Zeev. 1989. “Power, Capabilities, and Paradoxical Conflict Outcomes.” *World Politics* 41(2): 239-266.
Mason, David T. 1996. “Insurgency, Counterinsurgency, and the Rational Peasant.” *Public Choice* 86 (1/2): 63-83.
McDermott, Rose. 1998. *Risk-Taking in International Politics: Prospect Theory in American Foreign Policy*. Ann Arbor, MI: University of Michigan Press.
Mearsheimer, John J., and Stephen M. Walt. 2013. “Leaving Theory Behind: Why Simplistic Hypothesis Testing is Bad for International Relations.” *European Journal of International Relations* 19 (3): 427-457.
Melnik, Constantin. 1967. *The French Campaign Against the FLN*. Santa Monica, CA: RAND Corporation, https://www.rand.org/content/dam/rand/pubs/research_memoranda/2005/RM5449.pdf.

Mendelsohn, Jack and Walkling Sarah. 1996. “The CFE Treaty Review Conference: Strengthening the ‘Cornerstone’ of European Security.” *Arms Control Today* 26 (3): 3-6.
Miakinkov, Eugene. 2011. “The Agency of Force in Asymmetrical Warfare and Counterinsurgency: The Case of Chechnya.” *Journal of Strategic Studies* 34 (5): 647-680.
Middlebrook, Martin. 2006. *The First Day on the Somme*. Barnsley: Pen & Sword Books.
Miller, Sergio. 2012. “Malaya: The Myth of Hearts and Minds.” *Small Wars Journal*, April 16.
Millett, Allan R., Williamson Murray, and Kenneth H. Watman. 1986. “The Effectiveness of Military Operations.” *International Security* 11 (1): 37-71.
Mitzen, Jennifer and Randall L. Schweller. 2011. “Knowing the Unknown Unknows: Misplaced Certainty and the Onset of war” *Security Studies* 20 (1): 2-35.
Moorcraft, Paul. 2012. *Total Destruction of the Tamil Tigers: The Rare Victory of Sri Lanka’s Long War.* Barnsley: Pen and Sword Books Ltd.
Morgenthau, Hans J. 1948. *Politics Among Nations: The Struggle for Power and Peace*. New York: Alphred A. Knopf.
Morrow, James D. 1985. “A Continuous-Outcome Expected Utility Theory of War.” *Journal of Conflict Resolution*. 29 (3): 473-502.
Morrow, James D. 1987. “On the Theoretical Basis of a Measure of National Risk Attitudes.” *International Studies Quarterly* 31 (4): 423-438.
Morrow, James D. 1989. “Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining.” *American Journal of Political Science* 33 (4): 941-972.
Mueller, John E. 1980. “The search for the ‘breaking point’ in Vietnam: The statistics of a deadly quarrel.” *International Studies Quarterly* 24 (4): 497-519.
Murray, Williamson. 2011. *War, strategy, and military effectiveness*. Cambridge: Cambridge University Press.
Nagl, John A. 2002. *Learning to Eat Soup with a Knife: Counterinsurgency Lessons from Malaya and Vietnam*. Chicago: Chicago University Press.
Noonburg, Virginia W. 2014. *Ordinary Differential Equations: From Calculus to Dynamical Systems*. Washington DC: Mathematical Association of America.
Nowlin, Matthew C. 2011. “Theories of the Policy Process: State of the Research and the Emerging Trends.” *Policy Studies Journal* 39 (S1): 41-60.
O’Hanlon, Michael. 2003. “Estimating Casualties in a War to Overthrow Saddam.” *Orbis* 47 (1): 21-40.
Oliker, Olga. 2001. *Russia’s Chechen Wars 1994-2000.* Santa Monica, CA: RAND Corporation, https://www.rand.org/pubs/monograph_reports/MR1289.html.
O’Loughlin, John, Frank D.W. Witmer, Andrew M. Linke and Nancy Thorwardson. 2010. “Peering into the Fog of War: The Geography of the WikiLeaks Afghanistan War Logs, 2004-2009.” *Eurasian Geography and Economics* 51 (4): 472-495.

Organski A.F.K. and Jacek Kugler. 1980. *The War Ledger.* Chicago, IL: University of Chicago Press.
Page, Scott E. 2018. *The Model Thinker: What You Need to Know to Make Data Work for You*. New York: Basic Books.
Paul, Christopher, Colin P. Clarke, Beth Grill, and Molly Dunigan. 2013. *Paths to Victory: Lessons from Modern Insurgencies.* Santa Monica, CA: Rand Corporation.
Paul, Christopher, Colin P. Clarke, Beth Grill, and Molly Dunigan. 2016. "Moving Beyond Population-Centric vs. Enemy-Centric Counterinsurgency." *Small Wars and Insurgencies* 27 (6): 1019-1042.
Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
Pedroncini, Guy. 1990. *Journal de Marche de Joffre.* Vincennes: Service historique de l'armee de terre.
Perry, W., and J. Moffat. 1997. "Measuring the Effect of Knowledge in Military Campaigns." *Journal of the Operational Research Society (UK)* 48 (10): 965-972.
Peterson, Christopher, Nansook Park and Patrick J. Sweeney. 2008. Group Well-Being: Morale from a Positive Psychology Perspective." *Applied Psychology: An International Review* 57 (1): 19-36.
Philpott, William. 2006. "The Anglo–French Victory on the Somme." *Diplomacy and Statecraft* 17 (4), 731-751.
Pitcher T.J., and Wyche C.J. 1983. "Predator-Avoidance Behaviors of Sand-Eel Schools: Why Schools Seldom Split". In *Predators and Prey in Fishes. Developments in Environmental Biology of Fishes, vol 2*, edited by Noakes D.L.G., Lindquist D.G., Helfman G.S., Ward J.A., 193-204. Dordrecht: Springer.
Popplewell, Richard. 1995. "Lacking Intelligence: Some Reflections on Recent Approaches to British Counter-Insurgency 1900–1960." *Intelligence and National Security* 10 (2): 336–352.
Powell, Robert. 1987. "Crisis Bargaining, Escalation, and MAD." *American Political Science Review*. 81 (3): 717-736.
Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. New Jersey: Princeton University Press.
Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60 (1): 169-203.
Prior, Robin and Trevor Wilson. 2005. *The Somme. Updated edition*. New Haven: Yale University Press.
Reiter, Dan and Allan C. Stam. 1998. "Democracy, War Initiation, and Victory." *American Political Science Review* 92 (2): 377-389.
Reiter, Dan. 2003. "Exploring the Bargaining Model of War." *Prospective on Politics* 1 (1): 27-43.
Reiter, Dan, Allan C. Stam and Michael C. Horowitz. 2016. "A Deeper Look at Interstate War Data: Interstate War Data Version I.I." *Research and Politics* 3 (4): 1-3.

Renzi, Fred. “Networks: Terra Incognita and the Case for Ethnographic Intelligence.” *Military Review* 86 (5): 16-23.
Reza, Fazlollah M. 1961. *An Introduction to Information Theory*. New York, NY: McGraw-Hill Book Company.
Richardson, L.F. 1960. *Statistics of Deadly Quarrels*. Pittsburgh: Boxwood Press.
Rose, Elihu. 1982. “The Anatomy of Mutiny.” *Armed Forces and Society* 8 (4): 561-574.
Rosen, Steven. 1972. War Power and the Willingness to Suffer. In *Peace, War, and Numbers*, edited by Bruce M. Russett, 167-83. Beverly Hills, CA: SAGE Publications.
Rubinow, S.I. 2003. *Introduction to Mathematical Biology*. Mineola, NY: Dover Publications.
Russell, John. 2011. “Kadyrov’s Chechnya – Template, Test or Trouble for Russia’s Regional Policy.” *Europe-Asia Studies* 63 (3): 509-528.
Schelling, Thomas C. 1966. *Arms and Influence* New Haven: Yale University Press.
Schram, Peter. 2019. “Managing Insurgency.” *Journal of Conflict Resolution*, March 29, 1-35.
Sean M. Lynn-Jones. 1995. “Offense-Defense Theory and Its Critics,” *Security Studies* 4 (4): 66-491.
Sepp, Kalev I. 2005. “Best Practices in Counterinsurgency.” *Military Review* 85 (3): 8-12.
Shannon, Claude E. 1948. “A Mathematical Theory of Communication.” *The Bell System Technical Journal* 27 (3): 379-423.
Sheffield, Gary. 2015. *The Somme: A New History.* London: Cassell Military Paperbacks.
Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." In *Peace, War, and Numbers*, edited by Bruce Russett, 19-48. Beverly Hills, CA: SAGE Publications.
Singer, J. David. 1988. “Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985.” *International Interactions* 14 (2): 115-132.
Slantchev, Branislav L. 2003. “The Power to Hurt: Costly Conflict with Completely Informed States.” *American Political Science Review* 97 (1): 123-133.
Slantchev, Branislav L. 2011. *Military Threats: The Costs of Coercion and the Price of Peace*. Cambridge: Cambridge University Press.
Slantchev, Branislav L. and Ahmer Tarar. 2011. “Mutual Optimism as a Rationalist Explanation of War.” *American Journal of Political Science* 55(1): 135-148.
Šmíd, Tomáš, and Miroslav Mareš. 2015. “’Kadyrovtsy’: Russia’s Counterinsurgency Strategy and the Wars of Paramilitary Clans.” *Journal of Strategic Studies* 38 (5): 650-677.
Smith, Leonard V. 1995. “War and ‘Politics’: The French Army Mutinies of 1917.” *War in History* 2 (2): 180-201.
Smith, Niel A. and Nathan W. Toronto. 2010. “It’s All the Rage: Why Mechanization Doesn’t Explain COIN Outcomes” *Small Wars and Insurgencies* 21 (3): 519-528.

Speidel, Michael P. 2002. “A History of Indo-European ‘Mad Warriors’.” *Journal of World History* 13 (2): 253-290.
Stam, Allan C. 1996. *Win, Lose, or Draw: Domestic Politics and the Crucible of War*. Ann Arbor: University of Michigan Press.
Staniland, Paul. 2012. “Organizing Insurgency: Networks, Resources, and Rebellion in South Asia.” *International Security* 37 (1): 142-177.
Strogatz, Steven H. 2015. *Nonlinear Dynamics and Chaos. With Applications to Physics, Biology, Chemistry, and Engineering*. Boulder, CO: Westview Press.
Sullivan, Patricia L. 2007. “War Aims and War Outcomes: Why Powerful States Lose Limited Wars.” *Journal of Conflict Resolution* 51 (3): 496-524.
Teamey, Kyle and Jonathan Sweet. 2006. “Organizing Intelligence for Counterinsurgency.” *Military Review* 86 (5): 24-29.
Thomas, Timothy L. 2000. “Grozny 2000: Urban Combat Lessons Learned.” *Military Review* 80 (4): 50-58.
Toft, Monica Duffy, and Yuri M. Zhukov. 2012. “Denial and Punishment in the North Caucasus: Evaluating the Effectiveness of Coercive Counter-Insurgency.” *Journal of Peace Research* 49 (6): 785-800.
Toft, Monica Duffy, and Yuri M. Zhukov. 2015. “Islamists and Nationalists: Rebel Motivation and Counterinsurgency in Russia’s North Caucasus.” *American Political Science Review* 109 (2): 222-238.
Trenin, Dmitri. 2001. *The End of Eurasia: Russia on the Border Between Geopolitics and Globalization*. Washington, DC: Carnegie Endowment for International Peace.
Tse-tung, Mao. 1989. *On Guerrilla Warfare*. U.S. Marine Corps Reference Publication FMFRP 12-18, Washington, DC: Headquarters United States Marine Corps, https://www.marines.mil/Portals/1/Publications/FMFRP%2012-18%20%20Mao%20Tse-tung%20on%20Guerrilla%20Warfare.pdf.
Valentino, Benjamin, Paul Huth, and Dylan Balch-Lindsay. 2004. “Draining the Sea: Mass Killing and Guerilla Warfare.” *International Organization* 58 (2): 375-407.
Van Evera, Stephen. 1998. “Offense, Defense and the Causes of War.” *International Security* 22 (4): 5-43.
Wagner, Harrison R. 1994. “Peace, War, and the Balance of Power.” *American Political Science Review* 88 (3): 593-607.
Watts, Stephen, Jason H. Cambell, Patrick B. Johnston, Sameer Lalwani, Sarah H. Bana. 2014. *Countering Other’s Insurgencies: Understanding U.S. Small-Footprint Interventions in Local Context*. Santa Monica, CA: RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_reports/RR500/RR513/RAND_RR513.pdf.
Whitehead, Ralph J. 2013. *The Other Side of the Wire: Battle of the Somme with the German XIV Reserve Corps, 1 July 1916. Volume 2*. Solihull: Helion & Company.
White, John A. 1964. *The Diplomacy of the Russo-Japanese War*. Princeton, NJ: Princeton University Press.

Wing, Susanna D. 2016. “French Intervention in Mali: Strategic Alliances, Long-Term Regional Presence?” *Small Wars and Insurgencies* 27 (1): 59-80.
Wohlforth, William C. 1987. “The Perception of Power: Russia in the Pre-1914 Balance.” *World Politics* 39 (3): 353-381.
Wright, Quincy. 1942. *A Study of War. 2 Vols*. Chicago, IL: University of Chicago Press.
Yun, Robet K. 2003. *Case Study Research: Designs and Methods. Third Edition*. London: Sage Publications.
Zeytoonian, Dan, Laura Geldhof, Maureen Green, Remi Hajjar, Chris Litwhiler, Christine Locke, James Myers, David Perrine, Cameron Weathers, David Smith, and Christine Watson. 2006. “COIN Operations and Intelligence Collection and Analysis.” *Military Review* 86 (5): 30-37.
Zhukov, Yuri. 2008. “Examining the Authoritarian Model of Counter-insurgency: The Soviet Campaign Against the Ukrainian Insurgent Army.” *Small Wars and Insurgencies* 18 (3): 439-466.
Zhukov, Yuri. 2014. *“A Theory of Indiscriminate Violence.”* PhD diss., Harvard University