Predicting War: Assessing the Viability of the Offense-Defense Theory

van Sliedregt, Christiaan

Predicting War: Assessing the Viability of the Offense-Defense Theory

MAIR Global Conflict in the Modern Era

Faculty of Humanities

Master's thesis

Universiteit Leiden

Christiaan van Sliedregt

S2342391

Word Count: 14977

Table of Contents

List of abbreviations

APC	Armored Personnel Carrier
AT	Anti-Tank
CFM	Correlation of Forces and Means
ICBM	Inter Continental Ballistic Missile
IFV	Infantry Fighting Vehicle
MBFR	Mutual and Balanced Force Reduction
MBT	Main Battle Tank
NATO	North Atlantic Treaty Organization
ODT	Offense-Defense Theory
ODB	Offense-Defense Balance
POMCUS	Prepositioning Of Materiel Configured in Unit Sets
SLBM	Submarine Launched Ballistic Missile
SOF	Special Operations Forces
START	Strategic Arms Reduction Treaty
TNW	Tactical Nuclear Weapons

1. Introduction

At the time of writing, war has once again erupted in Europe on a scale few had imagined to ever occur again. In fact, it was often mentioned in academic debates concerning conflicts and war that the frequency of inter-state warfare was declining to such low numbers that many scholars deemed war between states as something from the past (Mandelbaum, 2010). Instead, some academics claim that the focus should lie on intra-state warfare which has become more frequent, with recent violent examples such as the Syrian and Libyan Civil Warfare (Kaldor, 2013). However, the recent escalation in Europe between Russia and Ukraine, supported by NATO, has brought back attention to areas where tensions are high such as the Nagorno Karabakh and the Taiwan Strait, and the necessity for explanatory models in order to answer the question why states go to war (Askarov, 2020).

Within the field of security studies and international relations, answering this question has always been an important focus of study (Brighton and Barkawi, 2011). The reason for this is obvious, if it would be possible to predict when and how a conflict would erupt, it would theoretically be possible to take steps to avoid it. So if it would be possible to identify the key factors that showed whether the chance of escalation is high or low, wouldn't it be possible to use this knowledge to prevent escalation? Many academics thought about this too and used this information to develop theories and models that could be used by policymakers and other key figures in governmental institutes to prevent tensions from escalating into conflicts. As such, terrible calamities such as war might be prevented, something that would surely benefit humankind (van Evera, 1998).

One of these theories is the Offense-Defense Theory (ODT) developed by Robert Jervis as a model to measure the intensity of a certain conflict area (Jervis, 1978). There are many proponents and opponents to the viability of this theory, however, one recurring point of criticism is that the theory hasn't been tested enough (Lynn-Jones, 1995). As such, this research will use the theory and apply it to the security dilemma that exists between the Russian Federation and NATO, in order to test the usability and the viability of this model and in doing so, answer whether this theory can explain when tensions flare and escalate, or why it cannot. For this purpose, the following research question will be used.

To what extent can the Offense-Defense Theory explain escalation and de-escalation between states looking at the case of relations between NATO and the Russian Federation/Warsaw Pact?

There are plenty of examples of regions and areas where strife and tensions continue to exist or re-emerge even after decades have passed such as the aforementioned Nagorno Karabakh Conflict or the Taiwan Strait but in short, this case has been selected best because of its abundance of sources.

To summarize, the research will focus on applying the model to two cases in two periods in time in the aforementioned region between the selected actors. We will measure what the situation was according to the theory and compare this to the relations between both groups and whether there is a relation between both. Ultimately we will have a case that should prove the theory is right, one that is assumed should prove it wrong to assure there is no bias by only choosing cases that could prove or disprove the theory. Yet the region and actors stay the same in order to not let any regional differences create diverging views on the theory. The use of the model will be explained later in the next chapter after which the case studies are examined followed by the comparison between both in order to answer the main research question.

2. Theoretical Framework

Before we dive deeper into the academic debate surrounding the Offense-Defense Theory it is important to understand the underlying fundamentals considering the security dilemma and its roots in Realist theories. The main fundament of Realist theories is that there is no rule of law governing states, instead, there is anarchy and states must take steps to assure their survival as other states will do the same (Shiping, 2009). As such, security, or lack of, plays an important role in this world where survival is key and competition is constant because to protect their possessions, states will often seek to control resources or territories outside their own, and the possibility of an eventual clash of interest increases radically. Therefore, to survive, a state must increase its capabilities to defend its territories and interests to protect itself against internal and external threats. However, since all states live in this perpetuated state of competition, it is often unclear what their intentions are and whether they are increasing their capabilities to defend interests or to expand and therefore pose a threat to others (Jervis, 1978).

This situation, in which it is not always possible to read each other's intentions correctly, is called the Security Dilemma. It is argued that this phenomenon, the Security Dilemma, is the leading source of conflict between states (Snyder, 1984). One state increases its security, for example through weapon procurements or technological developments, only to eventually decrease the security of another state as the result shifts the balance of power. This shift in the balance of power will not immediately create a dilemma, but as states are often unable to read each other's intentions they do not know whether these actions are taken in order to protect themselves or threaten others. ODT theorists believe the intensity of this situation can be measured and as such, help with explaining conflicts or even predicting possible areas of future conflicts in order to take steps to de-escalate in order to prevent them (Nilsson, 2012).

2.1 Offense-Defense Model Explained

The Offense-Defense Theory builds its fundament and plausibility on the assumption that any change in the balance of power in combination with the ability, or inability, to distinguish these actions as either being aggressive or defensive can increase the chance of conflict significantly. This doesn't sound implausible at all as several conflicts in the past have erupted precisely because of this occurrence (Shiping, 2010).

Robert Jervis established ODT as a way to measure and describe the intensity of this dynamic between states. This theory would be an important tool in order to determine the volatility of a certain conflict and the chances of it escalating into open war. In order to assess this, Jervis breaks this down into four scenarios. The first scenario has the highest likelihood of further escalating due to two factors, first, it is impossible to distinguish between defensive or offensive actions taken by each side, and secondly, offensive actions have an advantage over defensive actions. In the fourth scenario, there is little chance of escalation as both factors favor peaceful resolutions, actions are easily distinguishable from each other and overall defensive actions have an advantage over offensive actions, which for example would radically decrease the possibility of a successful surprise attack being launched against you (Jervis, 1978).

Graph 2.1 Offense Defense Theory

Offense-Defense theory	Offense has the advantage	Defense has the advantage
Offensive posture is not distinguishable from defensive one	1. Doubly dangerous	2. Security dilemma, but security requirements may be compatible
Offensive posture distinguishable from defensive one	3. No security dilemma, but aggression is possible. Warning given	4. Doubly stable

Source: Robert Jervis, *Cooperation Under the Security Dilemma 1978*

The hypothesis is that if the graph shows that the security situation is safe because the defense has the advantage and actions are distinguishable then the chances of tensions increasing and leading to escalation should be minimal. Whereas if the offense has the advantage and actions are not distinguishable events could lead to escalation and in many cases, this could be the outcome (Jervis,1978).

2.2 Offense-Defense Balance

The most important component is the Offense-Defense Balance (ODB) which is the tool to measure the situation to specify which of the above-mentioned scenarios can be attributed to it. In its essence, it identifies which of the key two factors, whether actions are distinguishable

and what has the advantage, offensive or defensive operations (Glaser and Kaufmann, 1998). However, academics that use the theory often differ in their approach. This is not entirely surprising since the theory did not offer a specific methodological approach when it was first presented. Generally speaking, it always involves the ODB in which the balance of power is measured between the different actors (Garfinkel and Dafoe, 2019).

3. Literature Review

This chapter can be divided into four sub-chapters that will discuss the most prevalent and recurring points of critique concerning ODT. The first chapter considers the misunderstanding critics often have with what the goal and purpose of the theory should be. The second chapter deals with the fact that many critics argue that the offense-defense balance can't be measured and therefore is not credible, the third chapter discusses the apparent over-reliance on the technology of the theory, and the last chapter focuses on whether the defense or offense can be distinguished from each other. The chapter concludes with a discussion on how all these issues and points of critique fit into the theory tests.

3.1 Fundamentals of the critique

Some authors argue that the foundations of the theory are underdeveloped and that it lacks an agreed definition which in turn results in the inconsistent application, which is commonly accepted as coming from a lack of testing. However, the most serious of their accusations concerning ODT is that it is flawed due to the inability to measure outcomes of wars, as they are so uncertain (Lynn-Jones, 1995).

Although indeed factual that outcomes of wars can rarely be predicted if at all, this is not the purpose of ODT. In other words, the criticism is often focused on the wrong aspects of the theory, since ODT is a tool used to predict the intensity of a certain conflict or situation and the chance of escalation. In short, it's not about predicting the outcome of the war but what the chances are the outcome is war (Biddle, 2001).

3.2 How credible is the model?

Another point of criticism is focused on the credibility of the model in the sense that it is capable of accurately measuring the security situation. In other words, if the offense-defense balance cannot be measured, the theory cannot predict or explain anything (Glaser and Kaufmann, 1998). Proponents on the other hand argue that critics overstate the difficulty of measuring the balance while others argue that estimates of the balance are sufficient for the model to be used accurately. As they argue, analytical tools capable of measuring the offense-defense balance are frequently being used in other fields and are the same used for military net assessment. These tools focus on force ratios between attackers and defenders and what means are necessary in order to force a breakthrough (Leiber, 2011). Within the offense-

defense theory, this is crucial since it will depend on whether offensive actions or defensive actions have the advantage.

For example, one of these tools is the force-to-space ratio devised by Lidell Hart, who concluded that greater force levels could significantly benefit the defender. This was based on the fact that during the initial investments the attacker's capability to exploit points on the battlefield increased as it can exploit gaps in areas where the defender provided little coverage. Beyond a certain investment level, the defender begins to saturate the battlefield, reducing any opportunities the attacker derives from differences in the two patterns of coverage. He argued that there was some force-to-space ratio necessary to provide sufficiently unbroken coverage of a front. If it was below this force-to-space ratio, the advantage would be to the attacker who could easily concentrate his forces on the weak points or gaps of the defensive line. This tool was extensively used by NATO in order to estimate the size of a defensive force in order to repel a Warsaw Pact invasion (Hart, 1960). As such, the tools have been used and tested extensively.

Critics argue that a more complete theory would include additional variables, such as power, measured in relative resources; and military skill which is a country's ability to effectively employ military technology. The main criticism concerning the credibility of the theory can therefore be summarized as a lack of components that give the necessary information in order to measure the balance of power (Kaufmann and Glaser, 1998). Proponents on the other hand argue that the precision the theory requires for measuring the balance of power is often lower than critics realize and that estimates are sufficient to provide the theory with substantial explanatory and predictive capabilities (Nilsson, 2012). This is again an example of the necessity to further test the theory in order to answer this question. Additional testing would produce results that could explain whether indeed the estimates provided by the theory are sufficient or whether more components are needed.

3.3 The concerns regarding its dimensions

Another point of criticism focuses on the reliance, or according to critics, overreliance on technological factors and the incapability to measure this accurately or give it any value. This is not surprising considering the complexity of predicting technological advancements in warfare situations (Shiping, 2010). This is mostly attributed to the case that military means are rarely ever used in their intended full-scale war scenarios. The reason for this is that war, thankfully, is a rare occurrence and as a result, it is impossible to fully assess how the system

will be used as it can never really be frequently experienced, tested, and improved as other systems (Leiber, 2000).

For example, European powers failed to predict the grueling attrition warfare of the First World War as they had failed to recognize the potential of new technologies, such as machine guns and chemical weapons, as they were used in outdated doctrines. And since they failed to predict this, the overall cost far outweighed the perceived benefits each side believed could be achieved (Biddle, 2001).

However, this is a good example of how technology should be measured, which is within the inclusion of doctrine. Military means should be addressed according to each side's doctrine since actors will use their means quite differently (Garfinkel and Dafoe, 2019). And as such, it's able to measure the technological factor and advantage or disadvantage this has over the other side. During the Second World War, for example, the German Empire already took advantage of numerous technological advancements concerning radios, improved motor vehicles, and other interwar inventions. This might explain the discrepancy between the smaller numbers of the forces of the Axis and an overall larger number of Allied forces during the first weeks of the invasions of France, Belgium, the Netherlands, and Norway and the eventual outcome in favor of Axis forces (Nilsson, 2012). As such, there is a possibility to measure the technological factor within the Offense-Defense Balance, yet the context in which it is perceived to be used and the effectiveness are key in this regard. But critics do seem to have a point when it comes to the inevitable complexity of it all.

The implications of geography are the least controversial of all the factors that are related to the ODB (Lynn-Jones, 1995). Terrain can have a strong effect on whether defense or offense has the advantage. Terrain that slows or channelizes movement generally puts a strain on logistics, therefore, strengthening the defense more than the offense. In a similar way distance to the objectives can affect the balance, the vaster the area necessary to traverse in order to reach the objectives the more this favors the defense (Garfinkel and Dafoe, 2019).

3.4 Concerns regarding distinguishability of actions

The last point of criticism is aimed at the inability to distinguish between offensive and defensive actions and weapons (Shiping, 2010). Critics argue that this is impossible since most, if not all, weapon systems can be used in multiple ways. A tank can be used to hold a defensive line or storm it, therefore it cannot be perceived as either defensive or offensive. Proponents of the theory argue that although weapons are indeed often multifunctional and do

not need to merely focus on one aspect, being either defensive or offensive operations, there are always situations in which certain specifications favor certain operations. A certain weapon system can be better suited for attacking than defending (Lieber, 2000 & Garfinkel and Dafoe, 2019).

Furthermore, the impact of a specific weapon or technology innovation and its performance properties should not be assessed in simple isolation but according to the impact it has on the state's ability to perform offensive or defensive actions (Lynn-Jones, 1995). Certain improvements in weapon systems are more favorable depending on the nature of the military operation. Improvements in mobility are generally perceived to be favoring offense while improvements in firepower favor defensive actions. The effects of innovations in protection, logistics, communication, and detection are more varied, depending on how specific innovations interact with force behavior (Gaser and Kaufmann, 1998). Distinguishability might therefore be easier to assess than most critics argue, especially when measured in combination with the technological factor and whether it is perceived to favor either a defensive or offensive strategy and usage.

3.5 Conclusion

In conclusion, the main point of critique was that the theory is too complex to be measured and cannot predict the outcome of a conflict. Yet the purpose of the theory is not to predict the outcome of a conflict but whether the chances of escalation are increasing or decreasing. War will not suddenly happen because one side has the advantage over the other, but certain events might escalate as the chances are higher or lower considering the balance. As such, the model could be used to determine which areas show higher chances of escalation and therefore the possibility of war as an outcome.

The second point focuses on the credibility of the theory as opponents argue that it is too detailed to be measured. However, tools were already used to accurately measure military force balances even before the theory was developed. Yet, what many on both sides agree on is that the theory might be incomplete and lack certain variables. In general, opponents argue that ballpark estimates are sufficient in order to come to a valid conclusion.

The third point focuses on the apparent over-reliance on the technology of weapon systems and that it's too complex to measure accurately. However, the focus should never be merely on the weapon system or technology but on how it fits in the context of military operations.

And the last point of critique focuses on whether actions can actually be distinguished from being offensive or defensive. This mainly comes from the fact that weapon systems can be used in multiple different ways and therefore be used both offensively and defensively. However, systems always favor one usage over the other in certain military operations. Therefore, these systems should always be reviewed within the context of how they're being used in military operations and the state's ability to use them effectively.

Overall, according to proponents the theory is often misused, is needlessly considered too complex and its variables are often measured without taking into consideration the context of the military operations in which they would be used. As such this theory test will assume that ballpark estimates are sufficient to come up with a viable analysis and take into consideration the military context and the summarized points of critique.

4. Methodology

4.1 Overview

In this chapter, we will discuss the methodology of the thesis and which sources will be used. The structure will therefore be the following; first, the selected method and the reason for its choice will be discussed. Second, the variables chosen from the theory will be explained, as their usage differs among scholars. Third, the case selection will follow, such as why these two cases have been selected, why this geographical location, the actors involved, the period in time, and finally why focus on these events. Lastly, the sources that have been selected will be examined which will be followed by a short summary of how each chapter is structured.

4.2 Case Method

The goal of this thesis is to test the viability and usability of the offense-defense balance and its model. In doing so, it would be possible to contribute to the current debate in which it is acknowledged that there is a lack of testing of the theory (Adams, 2004). In order to do that, we must apply this model to a selection of cases in which the results will be compared and analyzed in order to find out to what extent they can explain the reason why certain events almost led to escalation or indeed escalated or why not at all (Yin, 2018).

The following method will be employed through process tracing, we will examine whether indeed a certain event, in this case, military exercise, led or almost led to escalation, depending on how the situation was assessed in the graph (Yin, 2018). Process tracing is a method to assess whether there's a relation between cause and effect, or outcome. It is used to examine which mechanisms in place led to the outcome. In this study, the cause is the military exercise and the mechanisms in place are assessed through ODB leading to the aftermath, the outcome, which should show a correlation between the both (Yin, 2018).

For example, if the situation of the security dilemma can be classified as doubly dangerous it means the event was close to escalation or indeed escalated because tensions were extremely high. If. on the other hand, it can be classified as doubly stable, the result should be that the event was not considered threatening at all (Jervis, 1978). As such, each case goes through four phases; the first phase consists of a small outline and background of the military exercise. This is done in order to better understand the context of the military situation, something that was mentioned before in the literature review as the proponent's

object is often missing from case studies (Glaser and Kaufmann, 1998). The second part consists of analyzing and examining the various variables in order to determine how the situation can be assessed according to the model (Leiber, 2011). The next phase will be to determine where on the graph (Graph 4.1) the situation can be placed according to the variables. And finally, the direct aftermath of the event will be discussed and whether it correlates with the graph and therefore shows casualty with the graph (Yin, 2018).

By using comparative research, it is possible to focus on multiple outcomes of the model and analyze whether they differ and what is the reason for this differentiation. If we would only focus on one case it would be possible to overlook certain aspects that would only become visible after analyzing multiple cases. As such we will not make any adjustments to the model and use it in the same way in all selected cases (Yin, 2018).

Graph 4.1 Offense-Defense Theory

Offense-Defense theory	Offense has the advantage	Defense has the advantage
Offensive posture is not distinguishable from defensive one	1. Doubly dangerous	2. Security dilemma, but security requirements may be compatible
Offensive posture distinguishable from defensive one	3. No security dilemma, but aggression is possible. Warning given	4. Doubly stable

Source: Robert Jervis, *Cooperation Under the Security Dilemma 1978*

In order to remain unbiased and prevent a conclusion that has already been predetermined, it is necessary to use multiple cases when comparing, one to disprove the theory, and one to prove it. This will create the best circumstances in which the theory can be examined without the risk of being biased toward either outcome, whether the theory holds up or whether it fails. Due to the scope of this research, it is impossible to assess more than two cases with the in-depth analysis required for the model to be properly applied (Yin, 2018).

The application of ODT focuses for a large part on the model and its graph. As can be seen from the graph it consists of two main pillars which actions have the advantage, whether they be offensive or defensive, and whether these are distinguishable from each other (Jervis, 1978). As such, each case will be assessed according to these two pillars. The first part will consist of analyzing the different variables that need to be measured in order to ascertain

whether offense or defense had the advantage. (Lieber, 2011). While the second part focuses on whether actions are distinguishable from each other.

Since ballpark estimates should be sufficient in order for the theory to work, a binary system will be used that fits the graph, as it is only composed of four stages and two pillars. Instead of one or zero either high/low is used, to measure distinguishability, or favorable/unfavorable to measure whether the offense has the advantage. For example, if the majority of the variables show that distinguishability is low, it can be assessed that offensive and defensive posture is not distinguishable from each other. And the same goes for whether offense or defense has the advantage. If the majority of the variables show that offense is favorable compared to unfavorable it will be assessed as such. The last part is determining on the graph where the two axes meet (Jervis, 1978).

The result will be compared with historical records regarding the aftermath of the situation. For example, if the graph shows that the situation was doubly dangerous, historical accounts of the military exercise on both sides should show that there was an increased fear of escalation. This would show that there is indeed a casualty between the graph and events and as such, it can accurately predict when the chances of escalation are high and when not.

4.3 Measured variables

Offense or defense has the advantage

When using the theory, academics often differ in which components they investigate when measuring the balance of power. Some authors focus also on training and doctrine to establish whether the actions are distinguishable between offensive and defensive as a defensive doctrine will surely give the advantage to a defensive posture opposing an offensive one (Glaser and Kaufmann, 1998). However, generally speaking, there are a couple of components that are always used, being Force Size, Geography, and Technology. These components will be used to determine whether offensive or defensive actions have the advantage and will be measured according to the Favorable or Unfavorable binary system.

Force Size includes the military capabilities of each side, often measured in military assets. This will consist of three components, the first being Total Force Size, what each side has to its disposition, in order to assess which side has a military advantage in number and systems compared to the other side. Available Force Size includes military assets nearby the exercise area, and lastly, Involved Forces are what is used in the military exercise and how this affects the balance (Glaser and Kaufmann, 1998). When measuring the balance of forces,

it will be compared to the force ratio levels necessary for an offensive shown in Graph 4.2 (Hart, 1960). For example, as a rule of thumb, in order to attack a prepared opponent one needs to outnumber the enemy three to one. The assessment will be either Favorable or Unfavorable considering how the balance changes.

Graph 4.2 Force Ratio Model

Friendly Mission	Position	Friendly: Enemy
Delay		1:6
Defend	Prepared or fortified	1:3
Defend	Hasty	1:2.5
Attack	Prepared or fortified	3:1
Attack	Hasty	2.5:1
Counterattack	Flank	1:1

Source: McCarthy, Combat Values: *A Unified Input for Correlation of Forces and Means 2020*

The second variable focuses on Geography, which is also tied closely to Force Size as the location of the Available and Involved Forces is a geographical matter (Biddle, 2001). This can be categorized into three components; Force Disposition, which ties in with the latter one considering Force Size, Terrain, and Overall Operational Area. Force Disposition shows which side has the initial numeric advantage in the opening stages of a possible conflict this could shift the balance of power significantly. Regarding Terrain the analysis will focus on whether this favors the defender or attacker. Generally speaking, open areas favor mobility and thus the attacker, while urbanized and hilly terrain favor protection and thus the defender. The Overall Operation Area differs from the terrain as it will assess the broader geographical situation, such as, where reinforcements would arrive from and how fast (Hart, 1960). The classification will again be either Favorable or Unfavorable for all three components.

The last variable is Technology which consists of six areas or components; Mobility, Firepower, Protection, Logistics, Communication, and Detection (Glaser and Kaufmann 1998 & van Evera, 1998). The technological component is to determine whether these military capabilities favor an offensive or a defensive strategy and also to establish which side has a qualitative edge over the other (Lieber, 2000). Overall, it is considered that Mobility favors the offense and Firepower favors the defense. The other four are more varied depending on

the usage and military context. All six components will be classified as either Favorable or Unfavorable towards the offensive.

Posture is or isn't distinguishable from offensive or defensive

Lastly, whether it is possible to distinguish the actions of each side as either defensive or aggressive, is mostly determined by the rhetoric and actions states take towards each other. As such, the following variables, Trust, and Transparency will be used. Trust will be measured through treaties that were in effect, in progress, or even broken when it comes to arms control and force reductions. The more states trust each other, the more likely they look favorable to disarmament, while if they do not disarm they either still feel threatened or remain a threat. Trust and Transparency will be assessed as either being High if posture is distinguishable or Low, if not.

Transparency will be measured by examining how transparent each side was about its intended goals and actions. During the Cold War and later on, there were many diplomatic channels, some still in existence, between each side to mediate the chances of escalation through misunderstanding. Furthermore, there are many agreements considering military exercises between both sides dealing with necessary communications and information that needs to be provided to each side before the exercise starts. For example, the intended goal, the time period, and which forces are involved. Transparency depends on how binding and comprehensive these treaties were and whether they were adhered to.

4.4 Case Selection

As already mentioned before in the research question the location and actors of the cases have already been selected. This has been done for two reasons; the first is to narrow the scope of the cases to one particular region and involved actors as a broader scope would be unfeasible for this study. And secondly, to guarantee that the conditions of the security situation largely remain the same (Yin, 2018).

As such, the choice has been made to focus on Europe and the recurring tensions that exist on the borders between West-Europe and East-Europe. Therefore, the actors that will be examined are NATO and the Soviet Union, and later on Russian Federation (Cross, 2013). There are reasons why one can argue that the Russian Federation is a continuation of the Soviet Union, in this case, it's because it did not take long after the Cold War before its animosity towards NATO resurfaced and the dynamic returned. Therefore the Russian Federation will be considered a continuation of the Soviet Union where security interests are

involved. One example of this is the recurring claims of the Russian Federation that NATO expansion is hostile as the Soviet Union was guaranteed NATO would not expand eastwards (Baev, 2018).

The last question involved is why the choice was made to focus on military exercises as the events to be examined? The reason for this was the inherent characteristics of military exercises and their uniformity between cases. There's a form of military build-up, often near a border and some sort of simulated combat takes place often in order to gain experience, explore the effects of warfare, and or test strategies. Furthermore, military exercises are often a point of discussion as to what extent they lead to escalation (Kua and Blankenship, 2022). Not all tensions are uniform and it is hard to examine the commonality between for example the Cuban Missile of 1962 crisis and the 1960 U-2 Incident. It is however easier to focus on military exercises as their characteristics are largely the same (Yin, 2018).

Although many military exercises, were indeed only what they were supposed to be some exercises have been used as a ruse in order to invade and gain the initiative through a surprise attack. As such, tensions are therefore always higher between countries with strained relationships when military exercises are involved. The reason for this seems to be clear as any country would be cautious whenever large troop movements gather near its border (Huth, 1999). Another reason is that as the model measures whether actions are distinguishable from each other and if this is not the case, the chances for escalation during a military exercise are much higher as it might not be clear what the other side's intentions are (Snyder, 1984 & Jervis, 1978).

The first case is that of the Autumn Reforger 83 military exercise that could have resulted in open conflict between the two sides. This case should prove the theory as the situation was probably doubly dangerous at the moment considering the dynamics of the Cold War and as such, this military exercise came close to sparking World War Three (Heuser, 2018). It is, however, important to say that this recently came under discussion to what extent this could have sparked World War Three. At the moment, it is the case that comes closest to the expected hypothesis, the situation was Doubly Dangerous and the chances of escalation were high which was shown by the resulting tensions on both sides.

The second case that will be examined, is the military exercise Zapad 17. The intended goal of Zapad 17 was to defend the country from a large-scale attack coming from an adversary that showed similar characteristics as NATO (National Interest, 2017). It was feared by NATO to be a ruse for a possible aggressive move by the Russian Federation. This

did not happen and therefore, it is assumed that this event will disprove the theory. The situation was doubly dangerous after the annexation of Crimea.

4.5 Source selection

In order to measure Force Size the military capabilities of both sides will be examined through an analysis of sources provided by NATO and the International Institute for Strategic Studies, which annually publishes an assessment of the current military balance in the world. Furthermore, certain documentation will be examined, from historical archives, that shows which troops were involved during the exercise and nearby and which side had the local advantage. Considering this source, we can answer what the military balance was and how it changed during the exercise (Glaser and Kaufmann, 1998). Geographical sources will be used in order to determine the characteristics of the terrain. There are also multiple studies done by NATO and the RAND organization on possible conflict scenarios between NATO and the Russian Federation to assess the advantage of Force Dispositions and the Overall Operational Area Technology will be assessed by looking at certain assessments and documentation by military experts surrounding weapon systems in order to assess whether these favor the advantage or the offensive. This will be done in order to better understand the military context that seems to be often lacking in any application of the model. Although countries are of course reluctant to share technical data on their most modern weapon systems, it is possible to examine how these weapon systems would be used by looking at military doctrine and strategies (Garfinkel and Dafoe, 2019).

Furthermore, archival records and secondary sources will be used to analyze various treaties and events that could affect the level of Trust and Transparency between NATO and the Soviet Union. Agreements on force reduction and arms control will be examined to establish what level of Trust there was between both sides, such as START and NEW START in 2010. There are also multiple treaties that can be examined, the most prominent being the Helsinki Accords of 1975 and the Vienna Document of 2011 in order to measure Transparency.

4.6 Interpretation and structure

The study follows this strategy. By comparing various cases, selected to avoid any bias towards any explanation, it will be possible to analyze the validity and feasibility of ODT and its model. The results will show to what extent the model can explain any of the events that

happened and whether they indeed explain the aftermath that occurred.

The thesis is divided into two chapters each comprising a different case. Each chapter will begin with an introduction of the tensions that arose and the necessary context. Next, the various variables will be measured against the graph in order to eventually compare the results to the aftermath of the exercise. Finally, the conclusions of each case will be compared in order to come up with a comprehensive answer in order to reach final conclusions and add further recommendations.

5. Case analysis: Autumn Reforger 83

5.1 Introduction

The first case that will be examined is Autumn Reforger 1983 an annually recurring military exercise starting in 1967 by NATO. It was conceived to show the continued commitment to the defense of NATO and its capability to rapidly reinforce NATO units after US president Lyndon Johnson withdrew US troops from Europe. It consisted of a large-scale deployment of forces and often consisted of a division or more that would deploy to West Germany. It also served as the foundation of a plan to reinforce NATO presence in Europe, in which case it would chance into Operation Reforger (Heuser, 2018).

In essence, it involved troops redeploying from the USA to Europe where they would equip themselves from pre-positioned US military equipment stocks (POMCUS) in various places in Europe. As such these units could quickly become combat effective, without having to wait for heavy military equipment to arrive from overseas or by air where, in case of war, they would be vulnerable to Soviet attacks (Palmer, 2018).

However, it not only consisted of these drills but also of several war games and exercises such as Able Archer 83. Able Archer simulated command and control procedures, and how these transitioned from conventional through chemical and nuclear operations during wartime (Palmer, 2018). Able Archer 83 is generally believed to have come severely close to escalating into Nuclear war, although this is debated as to how close it really came to escalation (Heuser, 2018).

The structure is as follows, the Force Size, Geography, and Technology will be assessed then followed by Trust and Transparency in order to be summarized and compared to the aftermath of the exercise.

5.2 Force Size

As mentioned before, three components will be examined to measure their effect on the Offense-Defense Balance to distinguish whether offensive or defensive actions are favorable. Looking at the first component in the graph (Graph 5.1) shows the disparity between overall NATO forces and Warsaw Pact forces. It is visible that Pact forces outnumber NATO forces except for one component. However, Pact forces only hold a significant advantage over NATO in certain areas. This makes sense when taking into account the Soviet Doctrine that

relied heavily on the usage of mobile armored forces capable of performing a breakthrough while being supported by heavy use of artillery and fires (NATO Information Centre, 1980). NATO strategy relied on technological superiority in order to offset this advantage in numbers, this will be discussed later though (Miller, 1981). Yet, this does mean that Soviet forces hold a slight numerical advantage, however, meaning NATO forces lacked the necessary numerical advantage for an offensive. When we look at Nuclear Forces (Graph 5.2) the picture remains the same and the disparity between both sides is even less profound as Pact forces only hold a slight advantage over NATO forces regarding Nuclear forces (NATO Information Centre, 1984).

Graph 5.1 Total Force Numbers

Weapon Systems	NATO	PACT	Ratio -/-
Total Military	4.5 m	6.0 m	1/1,5
Divisions (Equivalents)	115	182	1/1.6
Main Battle Tanks (MBT's)	17.730	46.230	1/2.6
Anti-Tank (AT) Systems	19.170	36.400	1/1.9
Artillery and Mortars	14.700	36.600	1/2.5
Armored Personnel Carriers (APC's), Infantry Fighting Vehicle (IFV's) other Armoured Vehicles	39.580	94.800	1/2.4
Attack Helicopters	900	1.175	1/1.3
Transport/Support Helicopters	6.000	1.375	4.3/1
Total Aircraft	7.430	13.000	1/1.8

Source: NATO Information Centre, *Force Comparisons 1984*

Graph 5.2 Nuclear Forces

Types	NATO	PACT	Ratio -/-
Inter-Continental Ballistic Missile (ICBMs)	1040	1398	1/1.3
Submarine Launched Ballistic Missile (SLBMs)	632	945	1/1.4
Bombers	325	400	1/1.2
Tactical Nuclear Weapons (TNW)	1200	1600	1/1.3

Source: NATO Information Centre, *Force Comparisons 1984*

Graph 5.3 Available Force Number

Weapon systems	NATO	PACT	Ratio -/-
Total Military	1.5 m	2.0 m	1/1.3
Divisions (Equivalents)	38	61	1/1.6
MBT's	8.050	16.620	1/2
AT Systems	8.500	12.400	1/1.5
Artillery and Mortars	4.400	10.270	1/2.3
APC's, IFV's other Armoured	12.000	33.000	1/2.8
Attack Heli's	480	750	1/1.6
Transport/Support heli's	1.960	860	2.3/1
Total Aircraft in Europe	1.990	4.580	1/2.3

Source: NATO Information Centre, *Force Comparisons 1984*

Since the military exercise was focused around northern Europe, predominantly West-German forces near the vicinity of the theater of operations at the immediate outbreak of hostilities will be examined. The graph (graph 5.3) shows all available military assets for both sides that would be able to mobilize within 48 hours (NATO Information Centre, 1984). And as we can see the numerical advantage of Pact forces over NATO diminishes slightly. Overall Pact forces' numerical superiority remains in an overall ratio of two to one. Although

NATO's disadvantage in numbers diminishes slightly this is far from enough to achieve the necessary advantage of three to one required for an offensive to succeed (Hart, 1960).

Graph 5.4 Involved Force Numbers

Type	Amount
Troops	35.600
Armored Vehicles	2.050
Other vehicles	7.100

Source: Info Royal Netherlands Army, *Reforger '83 1983*

This year Autumn Reforger coincided with and included the Dutch military exercise, Atlantic Lion 83 a large Field Training Exercise that also involved a contingent of British troops and the troops involved in Autumn Reforger (Palmer, 2018). Although consisting of significant numbers only a few US and Dutch divisions were involved during Autumn Reforge alongside small elements of other militaries such as coming from the UK (Info Royal Netherlands Army, 1983). As such, the number of forces involved did not significantly shift the balance favorably towards NATO's side when it came to numbers.

To summarize, neither side held a substantial numerical advantage over the other. Warsaw Pact forces outnumbered NATO forces in some areas which would give them some advantages over NATO. However, NATO was sorely outnumbered by Pact forces and would have been unable to launch a conventional offensive that would require the necessary force advantages. Therefore we will classify this as Unfavorable towards offensive operations.

5.3 Geography

If we look at the geographical area we have to take into account three components, the first being the size of the potential Theater of Operations, Force Disposition, and how this could have affected a possible first strike advantage for NATO, and the Terrain of the area. And since all the exercises, were in or around West Germany they were mostly threatening to East Germany, as such, we can assume that this would be the potential Theater of Operations of any hypothetical conflict between both sides (NATO Information Centre, 1984).

Both sides relied heavily on the mobilization of forces to bring active-duty formations up to strength, however, when looking at the Theater of Operations it is clear that this is more favorable towards the Warsaw Pact than for NATO. The bulk of NATO's reinforcements had to be moved across the Atlantic and English channels. The Warsaw Pact on the other hand

could move many of its reserves over railways or roads. As such the advantage was more towards the Warsaw Pact (NATO Information Centre, 1984).

Force disposition remained still in favor of Warsaw Pact forces due to their numerical advantage over NATO forces. Although the involved NATO divisions could have made some territorial gains in a surprise attack, assuming Warsaw Pact forces would be caught off guard and required 48 hours to mobilize, overall speaking they would lack the necessary numbers to press on. As such, it can be assumed that any surprise attack would quickly be halted without any significant gains to show for it (Aspects of NATO, 1980).

Looking at the terrain of both East and West Germany consists of forests, hills, and urbanized terrain both favoring the defender as they provide ample cover and protection, yet also plenty of flat areas that would be suited for mobility, favoring the attacker. The Terrain was therefore not predominantly in favor of the defender or attacker. For example, efficient use of the terrain became an important cornerstone for NATO's defense as it built its defensive strategy on funneling PACT forces into areas where it could not efficiently employ its numerical superiority. Although it is not known whether Pact forces had similar plans for the defense of East Germany (Bluth, 2004).

Considering all examined components it is possible to deduce that the overall size of the operational theater was not in favor of NATO considering their dependency on reinforcements. However, looking at the disposition of forces on both sides, neither side had a considerable advantage over the other favoring the defense. Looking at the terrain it is harder to assess what is favored more, but considering the foundation of NATO strategy relied heavily on the terrain it is slightly more favorable towards the defense. In summarizing, all variables show that the defense was more favorable than the offense.

5.4 Technology

Considering technology, it is important to see this in the context of both sides' military doctrine. As mentioned earlier NATO relied on superior technology compared to Soviet and Warsaw Pact doctrine which favored quantity over quality. Warsaw Pact doctrine was therefore much more focused on the offense compared to NATO's more defensive doctrine.

During this time many weapon systems had already been employed or were being introduced that were highly mobile. Systems such as the Main Battle Tank (MBT), Infantry Fighting Vehicle (IFV), and various types of Helicopters and Fighter jets. More and more battlefield formations also became mechanized or motorized greatly increasing their mobility

(Miller, 1981). Mobility was therefore heavily favoring offensive operations.

There were also great leaps in Firepower as weapon calibers increased both in size and in the rate of fire. Also, more and more systems were equipped with higher caliber weapons such as troops transport which became IFVs. systems that could now also fight alongside the infantry and MBTs it supported (Miller, 1981). It also saw the increased introduction of various Anti-Tank Guided Missiles that saw such success during the Yom Kippur war of 1973.

Increases in Protection had generally been falling behind as the design of many weapon systems previously seemed to be focusing on increasing mobility and firepower. This is visible in the design of the Leopard-1, AMX-30, and T-62. Although this was starting to change with the introduction of heavier MBTs and IFVs such as the Leopard 2, M1 Abrams, and Challenger 1 (Miller, 1981).

When comparing Logistics both sides had fundamentally different approaches. Overall the speed and efficiency of logistical units had increased considerably compared to WW2, however, it was assumed that the pace of war would increase substantially too, meaning units would run out of supplies much quicker while needing resupply in order to remain combat effective. NATO forces relied on logistics and reinforcements in order to keep their formations combat effective. However, the Warsaw Pact relied on its numerical advantage, a formation would continue fighting till it was spent and then another formation would take over. As such there were vastly different perspectives on logistics on both sides (Miller, 1981).

The 1980s saw an increase in the availability of radio communication equipment in various weapon systems which favored the coordination between units. Overall speaking this could both favor the attacker and the defender. What was of substantial difference was how both sides employed their command structure. Soviet forces were commanded very centrally and its junior officers were not allowed to take any initiative. This fitted in their doctrine surrounding rehearsed set-piece attacks, favoring the offensive. NATO forces relied much more on their junior officers and their ability to take initiative and use the situation to their advantage. As such NATO forces were much more versatile compared to Warsaw Pact forces (NATO Information Service, 1984).

The ability to see the enemy before they can spot you has been crucial in any engagement. The developments concerning detection are perhaps key compared to all other components. The increase of detection systems tactically favored the defense as more advanced optics were installed on new weapon systems as well as night optics that became

available, although these were very rudimentary (Miller, 1981).

Considering the examined components, it can be assessed that Mobility and Protection were more favorable for offensive operations whereas Firepower was more favored by the defender. When it came to Logistics, Communication, and Detection these are neither favoring the offense nor defense more than the other. Generally speaking, NATO doctrine was more focused on the defense while the Warsaw Pact was more favored offensive operations. As such, technology was more favoring offensive operations than defensive.

5.5 Trust

In order to measure trust, the overall amount of arms treaties that are in effect and how the involved nations are following it is examined. And looking at 1983 trust on both sides can be considered Low. Even though this would change drastically mutual trust was still high in 1979 as the Soviet Union unilaterally withdrew 20.000 troops from East Germany as a sign of good faith towards NATO if they would refrain from positioning more nuclear weapons on European soil. This was all done in accordance with the Mutual and Balanced Force Reduction (MBFR) negotiations that had started five years earlier (Dean, 1983).

However, in 1983 progress was halted as the Soviet Union ceased its involvement in further talks. This was a reaction to the fact that the US had decided to deploy Pershing ICBMs on European soil (Arms Control Journal. 1983). the Soviets made a statement in which they warned that the deployment would lead them to reconsider their previous position taken in arms control treaties as they would have to take steps to counterbalance these missiles (Soviet Government Statement, 1983). According to the USA, this was a new capability in order to match existing Soviet Union nuclear capabilities. For the Soviet Union, this was seen as an escalation as these missiles could specifically target Soviet silos and command structures making a successful nuclear first strike possible. This was aggravated by the recently announced plans of the USA to develop a highly capable missile shield to counter Soviet missiles (Arms Control Journal, 1983). This greatly affected the negotiations and talks in Geneva considering arms control on nuclear weapons (Arms Control Journal, 1983).

5.6 Transparency

Examining the treaties that were in effect during the 1980's it can be concluded that Transparency was low. The main reason for this is that there were no binding rules in effect when it came to informing neighboring states of any planned military maneuvers. There was

some progress in changing this with the signing of the Helsinki Accords in 1975, which was an attempt to improve détente in Europe. In the document, all the parties involved agreed that efforts would be made to inform each other of any military maneuvers. There were some guidelines mentioned as to what this would entail such as inviting foreign observers to military exercises. However, the Helsinki Accords did not have the status of treaties and were not binding (Helsinki Accords, 1975). By the time of Autumn Reforger 83, relations between both parties had worsened because of several issues, including the deployment of Pershing missiles mentioned above. As such European states had stopped informing each other and were no longer offering visibility on each other's military maneuvers (Dean, 1982).

5.7 Summary

Summarizing the case will be done by looking at each of the examined variables. When looking at Technology this can be considered as being overall more favorable towards offensive actions than defensive, however, taking into consideration Force Size and Geography and the advantage go to the defender as both sides lack the necessary force advantage of 3:1 to overcome each other and the terrain and theater size is not in their advantage, especially for NATO. As it will be assessed that the defense has the advantage. When looking at whether offensive or defensive posture is distinguishable from each other this can be assessed as not being distinguishable as both Transparency and trust are considered Low. Looking at the graph the situation can be considered a Security Dilemma in which chances of escalation are present.

5.8 Aftermath

As mentioned before the military exercise Autumn Reforger 83 almost escalated into a nuclear exchange as the Warsaw Pact thought it was a ruse in order to take out Soviet nuclear retaliation capacity in a surprise first strike. Measures were taken and certain units were put on standby as the chance of escalation increased. The situation de-escalated because US leadership decided against putting their forces on increased standby as they wanted to wait until after the exercise was finished as they anticipated that it was related to the exercise. After the exercise finished Warsaw Pact forces gradually stepped down and tensions dwindled (Pry, 1999).

It is debated to what extent the crisis could have led to an escalation and full-scale nuclear exchange. Although it was generally acknowledged that the exercise caused the

reaction it is unclear to what extent the Soviets really considered the Exercise to be a ruse for a nuclear first strike by NATO (Heuser, 2018). Overall, it is assumed the Warsaw Pact was panicked by the fact that different procedures and actions were taken than normally were the case (Miles, 2020).

This does correlate with the result of the graph as the situation was not considered doubly dangerous, mostly because the offense did not have the advantage. The Soviets, however, became cautious as trust between both sides was degrading, which also correlates with the graph as mostly a lack of transparency led to increased tensions.

6. Case analysis: Zapad 17

6.1 Introduction

Zapad 2017 was a joint military exercise performed by the Russian Federation in conjunction with Belarusian forces in the month of September. It was a recurring exercise that continued after the dissolution of the Soviet Union. The scenarios of the exercise often simulated an attack on Russian territory or its allies to which Russian forces had to respond accordingly (National Interests, 2017). According to Russian authorities, the exercise was defensive in nature, yet this was heavily disputed by NATO as it often included offensive attacks on NATO territory such as in Zapad 2009 when a nuclear strike on Poland was simulated including offensive operations in and around Belarus. The scope of the exercise significantly changed after Zapad 99, the first one after the dissolution of the Soviet Union, showed that the Russian armed forces were incapable of repelling NATO forces leading to a revised defense concept (Dunin, n.d.).

Zapad 17 exercise scenario simulated the response to the threat of a large armed formation near Russian or Allied territory. It involved the rapid mobilization and redeployment of several Russian formations toward the Baltic Sea region and Belarussian exercise areas where they concluded several military drills. The involved units were formations from the 6th Air force and the 1st Tank Army from the Western military district and constituted around 13.000 troops and around 20 ships and support vessels (DW, 2017). The size of this exercise did not necessitate the need for mandatory formal notification to invite observers under the Vienna treaty.

Months before the exercise speculation began on how many troops would actually be involved and what could be the underlying motives other than those that had formally been disclosed. Experts were concerned that the exercise was a ruse to increase their military presence in Belarus near the Suwelki gap, threatening Poland and the Baltics (New York Times, 2017).

As with the first case, we will first examine whether offense or defense has the advantage by looking at the overall Force Size, Geographical situation, and Technological advantages or disadvantages after which we will examine whether posture was distinguishable or not by looking at Trust and Transparency resulting in the analyses of the aftermath and its relation to the assessment according to the ODT.

6.2 Force Size

When looking at the graph (Graph 6.1) there are certain small changes compared to the previous case that has to be explained before examining the graph. Since the Cold War, the leading military formation has changed from the division to the brigade or even battalion which is why divisions were not included in this graph. The necessity of anti-tank systems also changed considerably as the focus began to shift away from the tank-centric armies of the Cold War. Furthermore, missile technology advanced considerably and saw it being incorporated into many armored vehicles, such as IFVs (Hagler and Baker, 2018). As such the separate class of anti-tank systems and formations faded away as more weapon systems could successfully counter tanks. What is included though is the naval forces of each side due to the strong focus on the Baltic Sea Region and the inclusion of naval assets in the Zapad 17 exercise (DW, 2017). Furthermore, force multiplier aircraft have been included which is a class that consists of heavy UAVs, Airborne Early Warning (AEW), Intelligence Gathering (ISR), and Tanker planes. The reason for this is that after the cold war advancements in technology made these military assets much more important and their use more widespread (Bailey, 2013).

As the graph (Graph 6.2) shows the military situation had changed drastically compared to the balance of forces during the Cold War. The nuclear component shows that both sides had reached parity with each other as nuclear disarmament had significantly reduced the number of nuclear weapons both sides deployed. This is also understandable as technological advancements and the widespread understanding of mutually assured destruction led to the shrinking of nuclear stockpiles. For both sides, strategic nuclear deterrence became the main strategy (IISS Chapter Two, 2017).

Graph 6.2 Nuclear Weapons

Types	Russian Federation (RF)*	NATO	Ratio -/-
ICBMs	324	450	1.4/1
SLBM	13	14	1/1
Bombers	139	157	1.1/1

Source: IISS, Chapter Two, Comparative Defence Statisitcs, *The Military Balance 2017*

Examining conventional forces (Graph 6.1) the balance has shifted significantly in favor of NATO forces. In many cases, it employs a favorable force ratio, or even well above, for offensive operations. Only if the Russian Forces would mobilize their reserve forces it could hope to somewhat shift the balance in its favor. Especially in the aerial domain, an area NATO forces were already strong, NATO has significantly increased its advantage both numerically and technologically. As such the total force balance is to the advantage of NATO (IISS,017).

Graph 6.1 Total Force Size

Weapon Systems	RF	NATO	Ratio -/-
Active Military	0.8 m	3.2 m	4/1
Including Reserve Forces	3.1 m	5.5 m	1.7/1
MBT's	3465	9929	2,8/1
Artillery and Mortars	6242	26426	4/1
APCs, IFV's other Armored	16995	46717	2,7/1
Attack Heli's	344	1138	3,3/1
Transport/Support helicopters	393	4143	11/1
Tactical Aircraft	1101	5600	5/1
Transport Aircraft	440	1396	3/1
Force Multiplier Aircraft	76	1469	19/1
Aircraft Carriers and Amphibious Landing Ships	1	56	56/1
Cruisers, Destroyers, and Frigates	32	243	7,5/1
Tactical Submarines	49	132	2,7/1

Source: IISS, Chapter Two, Three, Four and Five, *The Military Balance 2017*

As the military exercise of Zapad 17 is focused on the eastern flank of NATO composing the Baltic states and Poland the focus will be on which forces in the area would be directly available for both sides in case of a sudden eruption of hostilities. And as can be examined in this case the balance of forces shifts dramatically in favor of the Russian Federation (Graph 6.3). Not only does the Russian Federation hold an advantage in numbers, but its formations,

are also heavier and better equipped, even its marines and airborne forces. NATO's numerical advantage in the air and on the sea is also diminished yet not as considerable compared to the ground forces. The military exercise, in its actual form, did not change this balance of forces as it was composed of elements of the abovementioned formations. Furthermore, it only consisted of around 13.000 troops, several brigades, and regiments, and 20 ships of the Baltic Sea Fleet (Dejevsky, 2017).

Overall speaking NATO has a considerable numerical advantage over the Russian Federation. Yet this changes when the focus is on local areas such as the Baltics in which case the balance of forces changes in favor of the Russian Federation. This means that in a short campaign it would have the advantage, however, the moment NATO mobilizes and increases its efforts to continue the war it is hard to imagine how the Russian Federation could withstand such numbers. As such this will be assessed as undecided since which side has an advantage is circumstantial to the military operation and objectives.

Graph 6.3 Available Forces

Unit Formations	RF	NATO	Ratio -/-
Armored Brigades/Regiments	6	2	3/1
Mechanized Brigades/Regiments	9	7	1.3/1
Artillery Brigades/Regiments	7	2	3.5/1
Airborne, Marine Brigades/Regiments	8	2	4/1
Light Infantry Brigades	0	4	0/4
SOF Brigades	3	0	3/0
Tactical Fighter Squadrons	14	8	1.75/1
Transport Plane Squadrons	2	2	1/1
Attack helicopter Squadrons	5	2	2.5/1
Transport helicopter Squadrons	4	3	1.3/1
Multiplier Squadrons	3	1	3/1
Submarines	2	5	1/2.5
Cruisers, Destroyers, Frigates	8	2	4/1

Source: IISS, Chapter Two, Three, Four and Five, *The Military Balance* 2017

6.3 Geography

In line with the first case, the geographical area consists of three components, the potential Theater of Operations, Force Disposition, and the Terrain. Considering the potential Theater of Operations the first issue for Russia is the fact that the Baltic Sea Region is enclosed by several NATO partners. Any of its reinforcing fleets would have to traverse this gap, one that would surely be heavily contested by NATO forces. The numerical advantage that the Baltic Fleet would have at the start of hostilities would quickly diminish as NATO reinforcements would be arriving quicker and unopposed compared to Russian naval reinforcements (Veebel and Sliwa, 2019).

However, looking at the situation on the ground the balance changes significantly due to the fact that Kaliningrad is almost entirely enclosing the Baltic states from Poland. NATO reinforcements would have to cross a narrow gap in order to reach the Baltics or go through Belarus, which probably would draw Belarus into the conflict leading to further escalation (Elak and Sliwa, 2016). Neither side can therefore be given the advantage, although Russian forces would have the advantage at the start of the conflict, later they would have a hard time dealing with NATO naval forces arriving in the theater.

The main advantage that the Russian Federation has comes from its favorable Force Disposition compared to NATO (IISS, 2016). In fact, the results of a wargame between NATO and Russian Federation forces published in 2016 showed that NATO forces at best could hold out for 72 hours before being overrun. The scenario involved a quick mobilization of only a portion of the available nearby Russian troops that would overwhelm NATO troops stationed in the Baltics. Russian troops held both a numerical advantage as well as an advantage in heavy weaponry which enabled them to overwhelm NATO defenders. NATO air forces, its biggest advantage, could not inflict enough losses on the Russians in order to halt their advance. As such the advantage was well on the side of the Russian Federation (Shlapak and Johnson, 2016).

As the study also showed, the Terrain was greatly in favor of the attacker especially if they consisted of highly mobile mechanized and armored formations. The area around the Baltics consists primarily of flat open terrain ideally for the attacker (Elak and Sliwa, 2016). The fact that NATO forces mostly consisted of light infantry formations only made this worse as these units perform badly in open terrain against more mobile and heavier equipped formations. And although the Baltics also consists of several large urban zones more suitable for the defender to halt an advance, they were easy to envelop and isolate posing no further

threats. Furthermore, the question remained whether the national governments of the Baltic states would accept the large-scale destruction of their capitals in order to win time for NATO reinforcements to arrive (Shlapak and Johnson, 2016).

6.4 Technology

At the end of the Cold War, many weapon programs were halted or adjusted and European states began to gradually decrease their armies and defense spending as the threat of war had ceased to exist (Hugh, 1998). Furthermore, NATO started to shift its focus on expeditionary missions and counter-insurgencies such as in Afghanistan during the War on Terror. The focus began to lie on developing light formations that could rapidly be deployed globally. NATO armies began to restructure their armies, putting less focus on developing their heavy formations, as they lacked funds to do both (Kaldor, 2013).

The Russian Armed Forces had fared badly in the conflicts after the Cold War. The disastrous results in Chechenia and the poor performance during the war in Georgia showed the necessity for the Russians to modernize. Even though new Russian weapon systems were developed and proudly presented, a lack of funds made large-scale implementation impossible. Instead, the Russians focused on modernizing their existing equipment as it was more cost-effective (Crane, Oliker and Nichiporuk, 2019).

Considering Mobility there had been a few advancements since the Cold War. Since armies were focused on counter-insurgency and rapid mobility tracks were exchanged for wheels on most military vehicles. Although missing the all-terrain capability of tracks, wheeled vehicles generally have higher mobility and were easier to maintain and therefore sustain on long operations. There was also a higher emphasis on strategic mobility, capable to deploy faster over the sea, air, and land routes (Hagler and Baker, 2018).

The advancements in Firepower were limited when it came to caliber size and rate of fire as neither increased significantly. Weapon calibers stayed relatively the same and the rate of fire only increased marginally for autocannons (Hagler and Baker, 2018). The changes included more competent fire control systems, precision-guided ammunitions, and as mentioned before missile technology. As a result of this, the accuracy and effectiveness of these weapon systems were increased. Furthermore, vehicles were often equipped with more armaments such as autocannons and missiles in order to counter multiple threats significantly increasing firepower (Bailey, 2013).

Advancements in Protection were considerable in both active defense and passive

defense. Sensors were used in order to disrupt or blind missiles, however, the effectiveness is debated. For ground vehicles, armor increased as both composite armor and applicable armor such as Explosive Reactive Armor became more effective and widespread, these were often applied to existing vehicles in order to extend their service time (Hagler and Baker, 2018). Stealth technology, in order to decrease radar signature and visibility, was introduced, although still only in limited numbers and mostly in Western countries (Bailey, 2013).

As mentioned before the biggest chance compared to weapon systems during the Cold War came from advancements in computing and networking (Bailey, 2013). This had the biggest effects on how weapon systems worked and the way armies operated. It became easier to coordinate between all branches of the military, for example between ground and air forces, greatly increasing the effectiveness of both (Hagler and Baker, 2018). As such Communication changed considerably compared to Cold War doctrine which is why the focus shifted from the Division being the center formation to the Battalion or Brigade (Crane, Oliker, and Nichiporuk, 2019).

Advancements in Detection were considerable and the usage of both night vision, thermal imaging, and other sensors became more widespread after its successful usage was demonstrated during the first Gulf War. Iraqi forces were at a disadvantage due to the fact that their vehicles lacked proper sensors in order to detect US forces in comparison the Coalition forces often did have these sensors on their vehicles (Hagler and Baker, 2018).

Compared to the Cold War the technological advancements of weapon systems despite being extensive were not visible. Existing weapon systems were upgraded with new modular armor, sensors, and networking capabilities and often with a focus on increased mobility (Bailey, 2013). Advancements in Firepower and Mobility favored the attacker while advancements in Protection favored the defender. The considerable advancements in Communication and Detection could favor any side but could give a tremendous advantage if the opposing side missed these capabilities. As such technology favored the offense considering NATO and Russian formations shared similar technologies and NATO's technological advancement had diminished.

6.5 Trust

In 2017 relations between NATO and the Russian Federation were very strained. Because of this Trust can be assessed as being Low when it comes to posture being distinguishable or not. The NEW START treaty was signed in 2010 in Prague further reducing NATO and Russia's nuclear arsenal (BBC, 2011). But although there were several Nuclear Arms Control Treaties

in effect this did not seem to increase Trust between both sides. One reason for this might be that the threat of nuclear warfare decreased significantly after the Cold War ended.

The main reason why the relations between both sides were strained, had everything to do with the situation in Ukraine. In 2014, only a few years ago before the exercise, the Russian Federation had forcefully invaded and annexed the Crimean Peninsula this was in violation of the Helsinki Accords of 1975. It was furthermore accused of being involved in the Ukrainian Civil War that erupted after the Maidan protests. Russia on the other hand criticized NATO for its unlawful expansion according to an agreement made by the heads of state of both the US and the Soviet Union after the reunification of Germany. This expansion was enveloping Russia and could only have as its goal to strain and contain Russia. As a result, both sides were skeptical and wary of each other's intentions (Kofman, 2017).

6.6 Transparency

Compared to Trust it is interesting to see that Transparency was actually on the complete opposite side and can be classified as High. The reason for this is the treaty that was signed in Vienna concerning military exercises in Europe called the Vienna Document. Compared to the Helsinki Accords that were mentioned in the first case, the Vienna Document was much more binding. The Helsinki Accords were more of a memorandum of understanding in which both parties promised to inform each other about military maneuvers, however, they were in no regard binding. This was radically different from the Vienna Document in which it is mandatory for states to inform each other of these military activities (OSCE, 2011).

Although both sides adhered to the rules stipulated in the treaty this could not have prevented a conflict over the exercise, probably fueled by the distrust already mentioned. Observers had not been invited formally according to the rules of the treaty. However, the involved amount of troops would not exceed the threshold in order to make the mandatory formal invitation necessary. This all seemed to be caused by the speculation that the exercise would involve more troops than Russia was presenting, and perhaps using the exercise as a ruse for aggression (Johnson, 2017). Observers were invited though and indeed confirmed that the number of troops involved never exceeded the amount Russia had presented (Dejevsky, 2017). As such, Transparency will still be considered as High as the treaty worked to increase transparency.

6.7 Summary

Just as with the first case we will look at the measured variables and see how they relate to each other in order to classify the situation according to ODT(Graph 2.1). While assessing force size we can see that there is a considerable difference between Total Force numbers and Available Force numbers. When looking at the Total Force number Russian Forces are significantly outmatched however when Available Forces are taken into account the situation chances the opposite way. In fact, it very much depends on the goal of the operation to which side is favored. However, both Geography and Technology can be classified as favoring the Offensive, and as such we will classify the Offense has the Advantage.

Classifying whether posture is distinguishable or not, is complicated due to the fact that although Trust is Low, Transparency can be considered high. Because Transparency was High and no binding Treaty was broken the advantage goes to posture being distinguishable between offensive and defensive. Comparing this to the graph shows that there is no Security Dilemma but Aggression is possible.

6.8 Aftermath

The Zapad 17 exercise led to considerable tension in Europe as military experts had speculated about Russia's true intentions behind the exercise and that the true amount of troops involved might be more than 100.000, which greatly exceeded the presented 13.000. Possible scenarios ranged from the annexation of Belarus to threatening the Suwelki Gap with an increased military presence furthering Russia's numerical advantage in the Baltics. As such, numerous requests were made towards Russia to include observers and the US committed troops to several ongoing exercises in the area (Johnson, 2017).

However, according to Western observers, the exercise never exceeded the amount that was presented, and as such the Vienna Document had never been violated. In hindsight, it seems that the results of the wargame of 2016, which showed that in case of a quick invasion Russia could occupy the Baltics within 72 hours, might have frightened military leaders and experts in Europe (Dejevsky, 2017 and Shlapak and Johnson, 2016). And as such, they were wearier of Russian military exercises fearing it could be used as a ruse to threaten the Baltics as Russia had shown it was more willing to use force after the successful annexation of the Crimea. This does correlate with the graph that even though Posture was distinguishable since Offense did have the Advantage military leaders were fearful of possible aggression even though it was considerably overreacted.

7. Summary of observations

In the previous two chapters, the cases have been separately examined. In this chapter, the outcomes and results of both cases will be compared before examining how they relate to the four points of criticism on the theory mentioned in the academic debate. The first point is, that it would be too complex to predict a conflict or a certain course of action to take place, second, it would be too complex to measure, third, that there was an over-reliance on technology, and last, if actions could be distinguishable.

The results of both cases showed promising outcomes and results when comparing the aftermath of the exercise with the result of the ODT. For example, looking at the first case, the aftermath of the exercise corresponds to the outcome of the graph. The reaction of the Warsaw Pact to the fairly routine exercise Autumn Reforger 83 makes sense when considering the situation was a Security Dilemma. Even though defensive actions had the advantage, the fact neither side could distinguish each other's actions as either being defensive or threatening meant that routine exercises such as Autumn Reforger almost led to escalation. In this case mostly because of the fact that distrust between the two sides had grown due to the stall in the progress on arms control treaties concerning force size and nuclear arms reductions.

Comparing the results for the Zapad 17 exercise, Western concerns over Russian military exercises are understandable, considering that the model shows Warnings are Given. This was mainly triggered by a significant imbalance between NATO forces and the Russian Federation forces putting NATO at a considerable disadvantage in a possible opening stage of a conflict. A more dangerous situation was averted mostly because of the fact that there was a high level of transparency considering military exercises on both sides. So overall speaking, comparing the outcome of the examination of the cases with the aftermath seems to show that the theory holds up.

Of course, it has to be said that both hypotheses were incorrect as the outcome of the model was not expected. It was expected that the situation during the Cold War was much more volatile and during 2017 was less. This could mean the theory was misapplied but the tools were properly used. In fact, in answering the four main critique points, the argument stays that the theory can be measured and is viable.

Taking the first point of criticism into account in both cases, it's the question of whether the aftermath could have been predicted by the theory, meaning that NATO would get concerned over Russian military exercises in the Baltic, or that the Soviet Union would

increase the readiness levels for its troops in response to routine exercises. The theory seems to be too narrow to give such a detailed prediction only focusing on how intense the situation is. However, as discussed in the literature review, the question is whether this is the point of the theory. Because in both cases the aftermath of the situation does fit with the outcome of the theory and as such holds up. It is not meant to predict the outcome of a conflict but whether the chances increase or not, which in both cases happened. And as such the theory could be used to explain reactions on certain events, such as military exercises, and how precarious the situation is. Therefore, the theory seems to be viable.

The second point of criticism was focused on whether the theory wasn't too complex to be measured. As mentioned before, there are tools available to measure the balance of forces. These models were used and indeed gave a detailed enough picture of the force balance of each side. However, this was predominantly the case when it concerned measuring the components of Force Size and Geography. Technology was harder to measure but this will be discussed in the next part. In fact, there are more models possible to use in order to give an even more detailed and comprehensive picture when it comes to Force Size and Geography. Unfortunately, these models are not accessible and are more specifically used within the military such as the Correlation of Forces Model (McCarthy, 2020). The measuring of components can be concluded to be less complicated than opponents seem to make it.

One point of criticism was directed against the overreliance on Technology for the theory to work. Technology would be hard to measure and therefore, its reliance on it would mean the overall downfall of the theory. Indeed, it seems safe to assume that mobility favors the offense and firepower the defense, it is not always clear to deduce what the other components favor, as it depends on the circumstances. That does not mean, however, that technology can't be measured, simply that it lacks tools that can be employed in the same way as measuring Geography and Force Size.

Technology could also be measured differently. In current cases, the overall technological level is taken into consideration as to whether weapon systems, in general, favor the offense or defense. For example, the invention of the machine gun, chemical gas, and other weaponry during the first world war shifted the technological level to favor defensive actions. Whether the inventions during the Second World War such as Tanks, Bombers, and Radio communications shifted this balance back in favor of the offensive (Leiber, 2011). However, the technological level could also be measured as the technological imbalance between both sides. A numerical advantage can be offset with a technological advantage such as NATO's strategy during the Cold War and the Autumn Reforger 83 case. However, this

could not be incorporated into the theory since proper tools are lacking, which seems to be in a disbalance.

Finally, there is the point of whether actions can be distinguishable or not. This was measured by including criteria of Trust and Transparency, where Trust is how much each side would trust each other's narrative on the actions taken, and Transparency is how visible those actions would be. Trust was harder to measure than Transparency mainly due to the fact that it is possible to examine each document and the rules that were agreed upon and whether these Treaties were followed or not. For example, the Vienna Document favors more visibility of each other's military actions than the Helsinki Accords.

Trust, however, was harder to measure compared to Transparency. Looking at Trust the focus was on Arms Control treaties but as was visible with the second case, although Arms Control treaties were in effect and being adhered to. The fact that Russia annexed Crimea was far more concerning to NATO compared to any other treaty that could have been broken.

8. Conclusion

This thesis revolved around testing the Offense-Defense Theory in order to assess and enhance its viability and usability of the theory. The main research question was as followed:

To what extent can the Offense-Defense Theory explain escalation and de-escalation between states looking at the case of relations between NATO and the Russian Federation/Warsaw Pact?

Looking back at the cases, it is safe to assume that the theory seems viable. The goal was to measure, through process tracing, whether there is a relation between the cause and outcome when using the mechanics of ODT. As mentioned in the previous chapter, in both cases the outcome of the situation, or in this case aftermath, is related to the mechanism shown in the model. If the situation is not a security dilemma but there is need for restraint because Warnings are Given (Graph 2.1) then military exercises can spark tensions. Because of this clear connection, between cause and outcome, the theory seems to be usable.

There were four main points of criticism identified that argued against the viability of the theory. The theory cannot predict the outcome of a war but this was never the point, the point was to predict whether tensions are high or not and as such conflicts could erupt. A strong case has been made in favor of the viability of the theory as the mechanisms of the theory show there is a relation between rising tensions and military exercises.

The focus was also on whether the theory would work with rough estimates as it was argued by some academics this would be too complex. The model is measurable as there are tools available to accurately assess the variables. It must be said though that not all variables are as easy to be measured, such as Trust and Technology. It is debatable to what extent technology can be measured.

However, regarding the overreliance of the theory on Technology, the question is how important technology is for the theory to work. As mentioned before, technology is indeed hard to measure, yet the other components are not. And with these components, an already substantial perspective on the balance of power can be given perhaps one substantial enough for the theory to work without including technology.

Yet this doesn't mean that the theory is proven, there is still more to be done. Adequate testing remains an important issue especially when it comes to distinguishability

and Technology, and it would help to create and improve tools in order to measure it. It is recommended to test the theory on different cases. Also, the focus was on military exercises and perhaps other events should be tried. Furthermore, it is the question of how important Technology is for the theory to work, furthermore, this could be negated by using models and tools that were inaccessible at the moment such as those reserved for the military.

Having said although that, the theory can be measured with rough estimates and simple tools the result is not specific either. However, it does show how high the chances of escalation are and that could be used to prevent tensions from spiraling into conflict. Such a model might be worth a lot in a time of recurring war.

9. Bibliography

Adams, K. (2004). Attack and Conquer? International Anarchy and the Offense-Defense-Deterrence Balance. *International Security*, *28*(3), 45-83.

Askerov, A. (2020). The Nagorno Karabakh Conflict. *Post-Soviet conflicts: The thirty years' crisis*, *55*.

Arms Control Journal (1983). Glimmers on arms control. *Nature*, *304*(5923), 197.

Arms Control Journal (1983). Poor outlook for arms control. *Nature*, *305*(5935), 561.

Arms Control journal (1983). Where now on arms control? *Nature*, *306*(5942), 411–412.

Baev, P. (2017). The Military Dimension of Russia's Connection with Europe. *European Security*, *27*(1), 82-97.

Bailey, J. (2013). *100 Military Inventions That Changed the World*. Skyhorse.

BBC News - US and Russian leaders hail nuclear arms treaty. (n.d.). BBC. Accessed on 5 May 2022, http://news.bbc.co.uk/2/hi/europe/8607985.stm.

Biddle, S. (2001). Rebuilding the Foundations of Offense-Defense Theory. *The Journal Of Politics*, *63*(3), 741-774.

Bluth, Christoph (2004). 'The Warsaw Pact and Military Security in Central Europe During the Cold War.' *Journal of Slavic Military Studies*, 17, 299-331.

Brauch, H. G., & Grimwood, T. (2014). *Jonathan Dean: Pioneer in Détente in Europe, Global Cooperative Security, Arms Control and Disarmament (SpringerBriefs on Pioneers in Science and Practice, 19)* (2014de editie). Springer.

Brighton, S and Barkawi, T (2011). Powers of War: Fighting, Knowledge,and Critique, *International Political Sociology*, 5, 126–143.

Crane, K., Oliker, O., & Nichiporuk, B. (2019). *Trends in Russia's Armed Forces: An Overview of Budgets and Capabilities*. RAND Corporation.

Cross, M. (2013). The Military Dimension of European Security: An Epistemic Community Approach. *Millennium: Journal Of International Studies*, *42*(1), 45-64.

David A. Shlapak and Michael W. Johnson (2016). Reinforcing Deterrence on NATO's Eastern Flank Wargaming the Defense of the Baltics. *RAND Corporation.*

Dejevsky, M. (2017, 22 September). *Russia was the target of Nato's own fake news*. The Independent. Accessed on 10 May 2022, https://www.independent.co.uk/voices/russia-zapad-2017-putin-west-fake-news-us-world-war-three-ukraine-belarus-nato-a7961856.html.

Deutsche Welle (www.dw.com). (n.d.). *Things to know about war games*. DW.COM. Accessed on 10 May 2022, https://www.dw.com/en/things-to-know-about-international-military-exercises/a-40468012.

Deutsche Welle (www.dw.com). (2017, 14 September). *What are Russia's Zapad war games?* DW.COM. Accessed on 10 May 2022, https://www.dw.com/en/what-are-russias-zapad-war-games/a-39702331.

Dunin, A. (z.d.). *News Article*. News Article – Center for Security Studies | ETH Zurich. Accessed on 10 May 2022, https://css.ethz.ch/en/services/digital-library/articles/article.html/109702

Evera, van, S. (1998). Offense, Defense, and the Causes of War. *International Security*, *22*(4), 5-43.

Garfinkel, B., & Dafoe, A. (2019). How does the offense-defense balance scale?. *Journal Of Strategic Studies*, *42*(6), 736-763.

Glaser, C., & Kaufmann, C. (1998). What is the Offense-Defense Balance and Can We Measure it?. *International Security*, *22*(4), 44.

Gordon, M. R., & Schmitt, E. (2017, 2 August). *Russia's Military Drills Near NATO Border Raise Fears of Aggression*. The New York Times. Accessed on 10 May 2022, https://www.nytimes.com/2017/07/31/world/europe/russia-military-exercise-zapad-west.html?hp&action=click&pgtype=Homepage&clickSource=story-heading&module=first-column-region&region=top-news&WT.nav=top-news

Hagler, G., & Baker, L. R. (2018). *The Evolution of Military Technology*. Van Haren Publishing.

Hart, Basil L., 'The Ratio of Troops to Space', *Royal United Services Institution Journal* 105/618 (1960), 201–12.

Huth, Paul (1999). Deterrence and International Conflict: Empirical Findings and Theoretical Debates. *Annual review of political science*, *2*, 25-48.

Jervis, Robert (1978). Cooperation Under the Security Dilemma. *World Politics*, *30* (2), 167-214.

Johnson, D. (2017, 14 december). *Zapad 2017 and Euro Atlantic Security*. NATO. Accessed on 12 May 2022, https://www.nato.int/docu/review/articles/2017/12/14/zapad-2017-and-euro-atlantic-security/index.html.

Kaldor, Mary (2013). 'In Defence of New Wars'. S*tability International Journal of Security and Development, 2* (1), 1-16.

Kuo, R., & Blankenship, B. D. (2021). Deterrence and Restraint: Do Joint Military Exercises Escalate Conflict? *Journal of Conflict Resolution*, *66*(1), 3–31.

Kofman, M., Migacheva, K., Nichiporuk, B., Radin, A., Tkacheva, O., & Oberholtzer, J. (2017). *Lessons from Russia's Operations in Crimea and Eastern Ukraine*. RAND Corporation.

Leszek Elak, Zdzisław Śliwa (2016) The Suwalki Gap- NATO's Fragile hot Spot. *Zeszyty Naukowe*, nr 2 (103*).*

Lieber, K (2011) Mission Impossible: Measuring the Offense-Defense Balance with Military Net Assessment, *Security Studies*, 20:3, 451-459.

Lieber, K. (2000). Grasping the Technological Peace: The Offense-Defense Balance and International Security. *International Security*, *25*(1), 71-104.

Lynn-Jones, Sean M. (1995), "Offense-Defense Theory and Its Critics," *Security Studies*, Vol. 4, No. 4.

Mandelbaum. M (2010). Is major war obsolete?. *Global Politics and Strategy*, Volume 40, 20-38.

McCarthy, M. J. E. (2020). Combat Values: A Unified Input for Correlation of Forces and Means. *Phalanx, 53*(4), 34–41.

McClintock, B., & Lilly, B. (2017, 12 September). *Zapad 2017: What It Reveals About the Prickly Russia-Belarus Relationship*. https://nationalinterest.org/feature/zapad-2017-what-it-reveals-about-the-prickly-russia-belarus-22271?nopaging=1.

Meyer Zum Felde, R. (2018). Beatrice Heuser/Tormod Heier/Guillaume Lasconjarias (Hrsg.): Military Exercises: Political Messaging and Strategic Impact. Rom: NATO Defence College, Februar 2018. *SIRIUS – Zeitschrift für Strategische Analysen*, *2*(3), 294–295.

Miles, Simon (August 1, 2020). "The War Scare That Wasn't: Able Archer 83 and the Myths of the Second Cold War". *Journal of Cold War Studies*, 22 (3): 86–118.

Miller, David (1981). *The Balance of Military Power: East v. West: An illustrated assessment comparing the weapons and capabilities of NATO and the Warsaw Pact*, London: Salamander Books.

Nilsson, M. (2012). Offense-Defense Balance, War Duration, and the Security Dilemma. *Journal Of Conflict Resolution*, *56*(3), 467-489.

Pry, Peter Vincent (1999). *War Scare: Russia and America on the Nuclear Brink*. Westport, Connecticut: Greenwood Publishing Group.

Rockoff, Hugh (1998). The Peace Dividend in Historical Perspective. *The American Economic Review*, *88* (2). 46-50.

Shiping, Tang (2009) The Security Dilemma: A Conceptual Analysis, *Security Studies*, 18:3, 587-623

Shiping, Tang (2010). Offence-defence Theory: Towards a Definitive Understanding. *The Chinese Journal Of International Politics*, *3*(2), 213-260.

Snyder, Glenn H. (1984). "The Security Dilemma in Alliance Politics". *World Politics*. 36 (4): 461–49.

Yin, R. (2018). *Case study research and applications* (6th ed.). SAGE Publications.

Veebel, V., & Sliwa, Z. (2019). The Suwalki Gap, Kaliningrad and Russia's Baltic Ambitions. *Scandinavian Journal of Military Studies*, *2*(1), 111–121.

Primary Sources

Helsinki Accords (1975). *The Final Act of the Conference on Security and Cooperation in Europe.* Accessed on 15 May 2022, https://web.archive.org/web/20160525015726/http://www1.umn.edu/humanrts/osce/basics/finact75.htm.

IISS (2017) Chapter Three: North America, *The Military Balance*, 117:1, 27-62,

IISS (2017) Chapter Two: Comparative defence statistics, *The Military Balance*,

117:1, 19-26,

IISS (2017) Chapter Four: Europe, *The Military Balance*, 117:1, 63-182,

IISS (2017) Chapter Five: Russia and Eurasia, *The Military Balance*, 117:1,

Info Royal Netherlands Army (1983). *Reforger '83 Brandy Station Aqua Marine Atlantic Lion*. Accessed on 8 May 2022, https://www.orbat85.nl/documents/Atlantic%20Lion%20%281983%29.pdf.

NATO Information Service, *Some Soviet and Warsaw Pact Military Facts and Figures 1980* (1980). Accessed on 8 May 2022, http://archives.nato.int/some-soviet-and-warsaw-pact-military-facts-and-figures.

NATO Information Service, *NATO and the Warsaw Pact: Force Comparisons* (Brussels, 1984). Accessed on 8 May 2022, https://archives.nato.int/nato-and-warsaw-pact-force-comparisons.

NATO Information Service, *Aspects of NATO Series 1 - n°1. Military posture. NATO and the Warsaw Pact* (1980). accessed on 27 April 2022, http://archives.nato.int/aspects-of-nato-series-1-n-1-military-posture-nato-and-warsaw-pact;isad?sf_culture=en.

OSCE (2011). Vienna Document 20110 On Confidence and Security Building Measures, accessed on 14 May 2022, https://www.osce.org/files/f/documents/a/4/86597.pdf.

Statement of the Soviet Union (27 May 1983). Nuclear Weapons and Arms Control, *Periodicals Archive Online* 25,5, pg. 232.