In Search of a Prime Mover:
An Investigation into the Relationship Between Technological Change and Shifts in the Fielded Military Capabilities of States

by Shahryar Pasandideh

B.A. in International Relations, May 2015, University of Toronto
M.A. in Political Science, August 2019, The George Washington University
M.Phil. in Political Science, January 2021, The George Washington University

A Dissertation submitted to

The Faculty of
The Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 31, 2024

Dissertation directed by

Charles L. Glaser
Professor of Political Science and International Affairs

The Columbian College of Arts and Sciences of The George Washington University certifies that Shahryar Pasandideh has passed the Final Examination for the degree of Doctor of Philosophy as of June 5, 2024. This is the final and approved form of the dissertation.

In Search of a Prime Mover: An Investigation into the Relationship Between Technological Change and Shifts in the Fielded Military Capabilities of States

Shahryar Pasandideh

Dissertation Research Committee:

Charles L. Glaser, Professor of Political Science and International Affairs, Dissertation Director

Alexander B. Downes, Professor of Political Science and International Affairs, Committee Member

Mike M. Mochizuki, Japan-U.S. Relations Associate Professor of Political Science and International Affairs, Committee Member

© Copyright 2024 by Shahryar Pasandideh
All rights reserved

Dedication

For my parents.

Acknowledgements

This dissertation project comes at the end of a long intellectual journey that I could not have completed without the support and guidance I received along the way. I thank Charles Glaser for offering an intellectual home in the Institute for Conflict and Security Studies (ISCS) at The George Washington University. ISCS provided a nurturing environment essential for my development as a scholar. I am grateful for the many stimulating discussions held with fellow graduate students and visiting predoctoral fellows. I am also grateful for the nurturing environment I encountered at the Belfer Center for Science and International Affairs at Harvard University during my two years as a visiting predoctoral fellow. My committee members, Alexander Downes, Charles Glaser, and Mike Mochizuki, endured much with grace over the years and I will forever be in their debt for the guidance and support offered over this long intellectual journey. To Charles Glaser, I owe an immense debt of gratitude in more ways than one.

I also want to thank the numerous colleagues who offered advice and support throughout this intellectual journey, among them Alexander Kirss, Alex Lin, Eleni Ekmektsioglou, Kendrick Kuo, and Stas Gorelick. Navigating life as an international student brought about numerous challenges and I am grateful for the support provided by departmental administrators and Ingrid Creppell in particular, who generously extended support when it was needed most while serving as Director of Graduate Studies. Generous financial support for this dissertation project was provided by The Simons Foundation, the Social Sciences and Humanities Research Council of Canada, and Canada's Department of National Defence. Although this dissertation project was undertaken as a graduate student at The George Washington University, I also wish to thank the many professors who offered

mentorship during my undergraduate years at the University of Toronto, among them James Reily, John English, Mari MacDonald, and Milena Methodieva.

Finally, I am grateful to my parents and my sister for their unwavering support. Thank you.

Abstract of Dissertation

In Search of a Prime Mover:
An Investigation into the Relationship Between Technological Change and Shifts in the Fielded Military Capabilities of States

Influential areas of international relations scholarship are predicated upon the notion that sizeable and abrupt shifts in the fielded military capabilities of states regularly transpire and are of consequence for international politics. Without what can be characterized as a prime mover to drive such shifts in the fielded military capabilities of states, variance in threat perceptions would lack a material basis. This dissertation project investigates whether technological change is a suitable candidate for the role of a prime mover driving sizeable and abrupt shifts in fielded military capabilities. In this dissertation, I develop an analytical framework that examines both monadic and dyadic mechanisms through which technological change can affect the fielded military capabilities of states. I then test whether the analytical framework accords with the empirical record through case studies focusing on advances in submarine technology in the ca. 1800-1945 timeframe and advances in uncrewed aircraft technology in the ca. 1910-2020 timeframe. I find that the effects of technological change tend to be considerably lagged and that states have considerable scope for adopting, countering, and offsetting new technologies to blunt if not negate the effects of technological change on their military standing. This dissertation's findings suggest that the field of international relations remains in search of a prime mover driving sizeable and abrupt shifts in fielded military capabilities at a time of growing interest in and concern about the implications of intensifying military-technological competition

Table of Contents

Dedication ..... **iv**

Acknowledgements ..... **v**

Abstract of Dissertation ..... **vii**

Table of Contents ..... **viii**

List of Figures ..... **xiii**

List of Tables ..... **xiv**

Chapter 1. Introduction **1**

Chapter 2. Shifts in Fielded Military Capabilities and the Necessity of a Prime Mover **17**

2.1. Candidates for a Prime Mover 36

Chapter 3. An Analytical Framework for Assessing the Military Effects of Technological Change **55**

3.1. Analytical Assumptions 56

3.2. Analytical Tenets 65

3.3. A Monadic Framework for Assessing the Pace and Magnitude of Shifts in Fielded Military Capabilities 71

3.4. A Dyadic Framework for Assessing the Pace and Magnitude of Shifts in Fielded Military Capabilities 91

3.5. Analytical Framework Conclusions 105

Part Two. Case Study: Advances in Submarine Technology, ca. 1800-1945

Chapter 4. Submarine Technology: Setting Up the Analysis **111**

Chapter 5. Advances in Submarine Technology ca. 1800-1900 118

5.1. Late-Nineteenth Century Technological Breakthroughs in Submarine Technology..... 134

Chapter 6. Advances in Submarine Technology ca. 1900-1914 144

6.1. Examining the Effects of Technological Change Through a Technology Development Lens..... 145

6.2. Advances in Submarine Technology..... 148

6.3. The Development of Countermeasures for Use Against Submarines 152

Chapter 7. Advances in Submarine Technology ca. 1914-1918 155

7.1. Wartime Advances in Submarine Technology..... 156

7.2. New Armaments for Submarines: A Primarily German Experience..... 159

7.3. New Coastal Submarines: A Primarily German Experience 165

7.4. Wartime Countermeasures Against Submarines 169

Chapter 8. Advances in Submarine Technology ca. 1918-1939 184

8.1. Advances In Submarine Armament..... 186

8.2. Advances In Submarine Sensing Technology..... 190

8.3. Advances In Communications Technology..... 192

8.4. Advances In Submarine Propulsion Technology..... 195

8.5. Experimentation and Conceptual Innovation in Submarine Design..... 200

8.6. Increased Competition in the Anti-Shipping Role..... 209

8.7. Developments in Anti-Submarine Countermeasures 213

8.8. Advances in Underwater Sensing Technology..... 220

Chapter 9. Advances in Submarine Technology ca. 1939-1945 223

9.1. Wartime Submarine Development 226

9.2. Wartime Measure-Countermeasure Dynamics 231

9.3. The Advent of "True" Submarines 252

Chapter 10. Reflecting Upon the Implications of Advances in Submarine Technology ca. 1800-1945 267

Part Three. Case Study: Advances in Uncrewed Aircraft Technology, ca. 1910-2020

Chapter 11. Uncrewed Aircraft Technology: Setting Up the Analysis 272

Chapter 12. Uncrewed Technology ca. 1910-1945 284

12.1. The First Uncrewed Fixed-Wing Aircraft 285

12.2. Technological Developments in the Interwar Period 290

12.3. Technological Developments During the Second World War 293

12.4. Assessing the Implications of Technological Change ca. 1900-1945 302

Chapter 13. Uncrewed Technology ca. 1945-1970 303

13.1. The Pursuit of Increasingly Automated Aircraft, Crewed and Uncrewed 304

13.2. The Development of Large Reusable Fixed-Wing Uncrewed Aircraft in the United States 319

13.3. Assessing the Implications of Technological Change ca. 1945-1970 327

Chapter 14. Uncrewed Technology ca. 1945-1970 330

14.1. Changing Requirements in the United States 330

14.2. Developments in the Soviet Union 333

14.3. The Enduring Limitations of LRFW Uncrewed Aircraft Technology 335

14.4. Background Technological Developments 336

14.5. The United States Renewed Pursuit of LRFW Uncrewed Aircraft in the 1970s 340

14.6. Israel's Development of LRFW Uncrewed Aircraft Technology in the 1970s and 1980s 346

14.7. Viewing Israeli Technological Developments in their Proper Context 349

14.8. The Gradual Adoption of Miniaturized LRFW Uncrewed Aircraft 353

14.9. The United States Renewed Pursuit of LRFW Uncrewed Aircraft in the 1980s 356

Chapter 15. Uncrewed Technology ca. 1990-2010 365

15.1. Changing Military Requirements in the Post-Cold War Era 367

15.2. Background Technological Developments 377

15.3. New LRFW Uncrewed Aircraft Designs 382

15.4. The Deficiencies of the Latest LRFW Uncrewed ISR Aircraft: The Case of the RQ-1 Predator 385

15.5. The Pursuit of More Capable LRFW Uncrewed ISR Aircraft 390

15.6. The Plateauing of LRFW Uncrewed Aircraft Technology 393

Chapter 16. Uncrewed Technology ca. 2010-2020 402

16.1. New Developments: LRFW Uncrewed Aircraft Weighing Under 1,000 Kilograms 403

16.2. New Developments: LRFW Uncrewed Aircraft Weighing Over 1,000 Kilograms 406

16.3. New Developments: Higher-End LRFW Uncrewed Aircraft 409

16.4. Renewed Efforts Toward Developing LRFW Uncrewed Strike Aircraft 418

16.5. The Observed Vulnerabilities of LRFW Uncrewed Aircraft to Countermeasures 426

Chapter 17. Reflecting Upon the Implications of Advances in Uncrewed Aircraft Technology ca. 1910-2020 ..... **431**

Conclusion. Reflecting Upon Relationship Between Technological Change and Shifts in the Fielded Military Capabilities of States **438**

Bibliography **449**

List of Figures

Figure 1. Shifts in the Fielded Military Capabilities of States and Security Dilemma Dynamics. 18

Figure 2. Gaps in International Relations Scholarship. 20

Figure 3. How States Can Expand Their Fielded Military Capabilities. 22

Figure 4. The Interaction of the Pace and Magnitude of Changes in Fielded Military Capabilities. 25

Figure 5. Fitting Adopting, Countering, and/or Offsetting into the Security Dilemma Framework. 33

List of Tables

Table 1. Adopt, Counter, and/or Offset: How States Can Respond to Technological Advances Pursued by Other States. 29

Chapter 1: Introduction

What are the *military effects* of technological change on international politics? Specifically, what is the relationship between technological change and shifts in the *fielded military capabilities* of states? Although international relations scholarship devotes limited direct attention to military technology and advances in military technology,[1] a significant portion of the literature requires *something* to drive shifts in the fielded military capabilities of states and is predicated upon the notion that such shifts in fielded military capabilities take place and are of consequence for international politics. Without what can be characterized as a prime mover to drive (be the driving force of)

[1] Although only a small subset of international relations scholarship directly examines military technology, there is a sizeable literature in which international relations scholarship more narrowly focuses on nuclear weapons, on issues concerning artificial intelligence and cybersecurity, as well as a sizeable literature with a broader focus on technological change that may touch upon certain areas of military technology. Consider, for example, Robert Gilpin, *War and Change in World Politics* (Cambridge, UK: Cambridge University Press, 1981); Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Ithaca, NY: Cornell University Press, 1989); Daniel Drezner, "Technological Change and International Relations," *International Relations* 33, No. 2 (2019): 286-303.; Jeffrey Ding and Allan Dafoe, "Engines of Power: Electricity, AI, and General-Purpose Military Transformations," *European Journal of International Security* 8:3 (2023): 377-397.; Daniel R. McCarthy, *Power, Information Technology, and International Relations Theory: The Power and Politics of US Foreign Policy and Internet* (London, UK: Palgrave Macmillan, 2015); Robert Paarlberg, "Knowledge as Power: Science, Military Dominance, and U.S. Security," *International Security* 29:1 (2004): 122-151.; Jon R. Lindsay, *Information Technology and Military Power* (Ithaca, NY: Cornell University Press, 2020); Michael Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, No. 2 (2018): 36-57.

International relations scholars also regularly draw upon the research of historians and other areas of academia that examine advances in military technology. Consider, for example, Martin van Creveld, *Technology and War: From 2000 B.C. to the Present* (New York: Free Press, 1989); Daniel R. Headrick, *The Invisible Weapon: Telecommunications and International Politics, 1851-1945* (Oxford, UK: Oxford University Press, 1991); Barton C. Hacker, "Military Institutions, Weapons, and Social Change: Toward a New History of Military Technology," *Technology and Culture* 35, No. 4 (1994): 425-440.; Max Boot, *War Made New: Technology, Warfare, and the Course of History, 1500 to Today* (New York: Penguin, 2006); William H. McNeil, *The Pursuit of Power: Technology, Armed Force, and Society Since A.D. 1000.* (Chicago, IL: University of Chicago Press, 1982).; Jeremy Black, *War and Technology* (Bloomington, IN: Indiana University Press, 2013).; Bernard Brodie and Fawn M. Brodie, *From Crossbow to H-Bomb: The Evolution of the Weapons and Tactics of Warfare* (Bloomington, IN: Indiana University Press, 1973).; Trevor N. Dupuy, *The Evolution of Weapons and Warfare* (Boston, MA: Da Capo Press, 1984).; George Raudzens, "War-Winning Weapons: The Measurement of Technological Determinism in Military History," *The Journal of Military History* 54, No. 4 (1990): 403-434.; Alex Roland, "Science, Technology, and War," *Technology and Culture* 36, No. 2 (1995), 83-100.; Kenneth Macksey, *Technology in War: The Impact of Science on Weapon Development and Modern Battle* (New York: Simon and Shuster, 1986).; Vernon W. Ruttan, *Is War Necessary for Economic Growth? Military Procurement and Technology Development* (Oxford, UK: Oxford University Press, 2006).

such shifts in the fielded military capabilities of states, variance – temporal and otherwise – in threat perceptions would lack a material basis and this would undermine the explanatory power of influential segments of international relations scholarship. This dissertation project investigates whether technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts – gains – in the fielded military capabilities of states.

International relations scholars have long recognized that technological change can affect the military power of states and, as such, affect international politics. The Cold War, which provided the backdrop for the formative years of the field, not only provided a litany of technological developments for scholars to consider but also incentivized the re-examination of the historical record for insights about the implications of previous experiences with technological change. Unsurprisingly, the advent of nuclear weapons – and closely associated non-nuclear military technologies such as ballistic missiles – had a major influence on the trajectory of the field, and many scholarly writings on military technology written in the Cold War either focused exclusively on developments involving nuclear weapons or compared technological developments in the nuclear era with those of prior historical experience.[2] Scholars explored parallels in the effects of the development and diffusion of railroads on military strategy in the years leading up to the First World War.[3] Many scholars also took a particular interest in developments in naval technology from circa 1850 to 1940 as this timeframe included multiple episodes of arms racing. While the naval arms race between Germany and the United Kingdom in the early

[2] See, for example, Kier A. Lieber, *War and the Engineers: The Primacy of Politics Over Technology* (Ithaca, NY: Cornell University Press, 2005).
[3] For example, Scott Sagan, "1914 Revisited: Allies, Offense, and Instability," *International Security* 11, No. 2 (1986): 151-175.

twentieth century offered a potential parallel to the nuclear arms race between the Soviet Union and the United States and highlighted the risks associated with arms racing, the ultimately unsuccessful efforts toward (naval) arms control in the Interwar Period offered a template for assessing the prospects for arms control in the nuclear era.[4] Although the dynamics of arms racing and arms control offered fertile grazing for international relations scholars of many stripes, two other literatures more closely examined historical experiences with technological change and the resulting research output has had a strong influence on international relations scholarship and scholarly views on the military implications of technological change.

Decades after its conclusion, burgeoning historical research on the First World War offered the budding field of international relations much to work with from a case through which to assess the effects of economic interdependence, a case through which to assess wartime constraints on the use of new military technologies (whether chemical weapons against combatants or submarines and aircraft against non-combatants)[5] and, more importantly for the present purposes, it also (re)introduced international relations scholarship to the conduct of the First World War and pre-war expectations as to how a prospective war would transpire.[6] To explain what hindsight suggested were erroneous expectations of future combat and, more generally, the conditions under which combatants operated on the Western Front in particular, international relations scholars

[4] For example, Charles L. Glaser, "When Are Arms Races Dangerous? Rational Versus Suboptimal Arming," *International Security* 28, No. 4 (2004): 44-84. On arms racing, see also Barry Buzan and Eric Herring, The Arms Dynamic in World Politics (Boulder, CO: Lynne Rienner Publishers, 1998)
[5] Consider, for example, Alexander B. Downes, *Targeting Civilians in War* (Ithaca, NY: Cornell University Press, 2008); Erik Gartzke and Yonatan Lupu, "Trading on Perceptions: Why World War I Was Not a Failure of Economic Interdependence," *International Security* 36:4 (2012): 115-150.
[6] For example, Stephen Van Evera, "The Cult of the Offensive and the Origins of the First World War," *International Security* 9, No. 1 (1984): 58-107.; Jack Snyder, "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984," *International Security* 9, No. 1 (1984): 108-146.

began to examine the causes of military stalemate and, more specifically, the determinants of whether offense or defense had an advantage. By comparing developments in warfare in the years leading up to the First World War with developments in the years leading up to the Second World War and, no less, contrasting the exceptionally variant experiences in those conflicts in terms of ground combat operations along the German-French-Belgian frontier region, scholars highlighted the role of qualitative technological changes whether advances in firepower in the lead up to 1914 or advances in armored combat vehicle technology in the lead up to 1939-1940.[7] The literature on offense-defense theory and, more generally, the notion of an offense-defense balance,[8] developed from this area of inquiry and thereby constituted the primary lenses outside of discussions involving nuclear weapons through which international relations scholars examined the implications of technological change in general and advances in military technology in particular.

[7] See, for example, Kier A. Lieber, *War and the Engineers: The Primacy of Politics Over Technology* (Ithaca, NY: Cornell University Press, 2005).

[8] Charles L. Glaser and Chaim Kaufmann, "What is the Offense-Defense Balance and Can We Measure It?" *International Security* 22, No. 4 (1998): 44-82.; Keir A. Lieber, "Grasping the Technological Peace: The Offense-Defense Balance and International Security," *International Security* 25, No. 1 (2000): 71-104.; Jack S. Levy, "The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis," *International Studies Quarterly* 28, No. 2 (1984): 219-238.; Sean M. Lynn-Jones, "Offense-Defense Theory and Its Critics," *Security Studies* 4, no. 4 (1995): 660-694.; Charles L. Glaser, *Rational Theory of International Politics* (Princeton, NJ: Princeton University Press, 2011).; Stephen Van Evera, *Causes of War: Power and the Roots of Conflict* (Ithaca, NY: Cornell University Press, 1999).; George Quester, *Offense and Defense in the International System* (New York: John Wiley Publishers, 1977).; Stephen Van Evera, "Offense, Defense, and the Causes of War," *International Security* 22, No. 4 (1998): 5-43.; Stephen Biddle, "Rebuilding the Foundations of Offense-Defense Theory," *Journal of Politics* 63, No. 3 (2001): 741-774.; Karen Ruth Adams, "Attack and Conquer? International Anarchy and the Offense-Defense-Deterrence Balance," *International Security* 28, No. 3 (2003/2004): 45-83.; Jack S. Levy, "The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis," *International Studies Quarterly* 28, No. 2 (1984): 219-238.; David Blagden, "When Does Competition Become Conflict? Technology, Geography, and the Offense-Defense Balance," *Journal of Global Security Studies* 6, No. 4 (2021): 1-23.

Attesting to its influence on international relations scholarship, offense-defense theory and the more general notion of an offense-defense balance permeated the literatures on the causes of war, escalation dynamics, alliances, and, more generally, the effects of shifts in military power on international politics. More recent scholarship has applied the core framework of an offense-defense balance to examine technological developments in the late twentieth and early twenty-first century in areas including naval warfare and cyberspace, thereby furthering the long-standing influence of an area of scholarly inquiry that rose to prominence early in the Cold War.[9] Writings on offense-defense dynamics, however, typically focus on the implications of technological change, broadly construed, rather than its causes, and therefore tend not to account for variance in its implications for international politics across time and space or the intricacies of technological development. The literature on military innovations, among others, worked to fill this gap by offering explanations for variation in both how states adopt what tend to be, in essence, technological innovations, and how states employ new technologies with or without doctrinal reconceptualization or adaptation.[10] Although many contributors to

[9] See, for example, Ben Garfinkel and Allan Dafoe, "How Does the Offense-Defense Balance Scale?" *Journal of Strategic Studies* 42, No. 6 (2019): 864-887.; Rebecca Slayton, "What is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41, No. 3 (2017): 72-109.; Erik Gartzke and Jon R. Lindsay, "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace," *Security Studies* 24, No. 2 (2015): 316-348.; Jon R. Lindsay, "Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence against Cyber Attack," *Journal of Cybersecurity* 1, No. 1 (2015): 53–67. For research extending this area of scholarship beyond cybersecurity, see Jon R. Lindsay, "Demystifying the Quantum Threat: Infrastructure, Institutions, and Intelligence Advantage," *Security Studies* 29, No. 2 (2020): 335–61.; Cameron L. Tracy and David Wright, "Modelling the Performance of Hypersonic Boost-Glide Missiles," *Science & Global Security* 28, No. 3 (2020): 135-170.; Antonio Calcara, Andrea Gilli, Mauro Gilli, and Ivan Zaccagini, "Will the Drone Always Get Through? Offensive Myths and Defensive Realities," *Security Studies* 31, No. 5 (2022): 791-825.; Stephen Biddle and Ivan Oelrich, "Future Warfare in the Western Pacific: Chinese Anti-access/Area Denial, U.S. AirSea Battle, and Command of the Commons in East Asia," *International Security*, Vol. 41, No. 1 (2016): 7-48.

[10] Barry R. Posen, *The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars* (Ithaca, NY: Cornell University Press, 1984).; Stephen P. Rosen, *Winning the Next War: Innovation and the Modern Military* (Ithaca, NY: Cornell University Press, 1991).; Michael C. Horowitz and Shira

the military innovations literature avowedly focus primarily on matters of military doctrine as opposed to new technologies, the vast majority of cases examined in the literature are characterized, if not defined, by one or more technological developments. Moreover, the military innovations literature tends to offer no explanations as to how or

---

Pindyck, "What is a Military Innovation and Why It Matters," *Journal of Strategic Studies* 46, No. 1 (2023): 85-114.; Matthew Evangelista, *Innovation and the Arms Race: How the United States and the Soviet Union Develop New Military Technologies* (Ithaca, NY: Cornell University Press, 1988).; Emily O. Goldman and Richard B. Andres, "Systemic Effects of Military Innovation and Diffusion," *Security Studies* 8, No. 4 (1999): 79-125.; Adam Grissom, "The Future of Military Innovation Studies," *Journal of Strategic Studies* 29, No. 5 (2006): 905-934.; Leslie C. Eliason and Emily O. Goldman, Editors, *The Diffusion of Military Technology and Ideas* (Palo Alto; CA: Stanford University Press, 2003); Michael C. Horowitz, *The Diffusion of Military Power: Causes and Consequences for International Politics* (Princeton, NJ: Princeton University Press, 2010).; Elizabeth Kier, *Imagining War: French and British Military Doctrine Between the Wars* (Princeton, NJ: Princeton University Press, 1997).; Kimberly Marten Zisk, *Engaging the Enemy: Organization Theory and Soviet Military Innovation, 1955-1991* (Princeton, N.J.: Princeton University Press, 1993).; Chris C. Demchak, *Military Organizations, Complex Machines* (Ithaca, N.Y.: Cornell University Press, 1991).; Dima Adamsky. *The Culture of Military Innovation: The Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the US, and Israel* (Stanford, CA: Stanford University Press, 2010).; Kendrick Kuo, "Military Magic: The Promise and Peril of Military Innovation," (Ph.D. dissertation, Political Science, The George Washington University, 2021); Benjamin Jensen, *Forging the Sword: Doctrinal Change in the U.S. Army* (Stanford, CA: Stanford University Press, 2016).; Stuart Griffin, "Military Innovation Studies: Multidisciplinary or Lacking Discipline?" *Journal of Strategic Studies* 40, No. 1–2 (2017): 196–224.; Williamson Murray and Allan R. Millett, Editors, *Military Innovation in the Interwar Period* (Cambridge, UK: Cambridge University Press, 2009); Owen Cote, "The Politics of Innovative Military Doctrine: The U.S. Navy and Fleet Ballistic Missiles" (Ph.D. dissertation, Political Science, Massachusetts Institute of Technology, 1996); Theo Farrell and Terry Terriff, Editors, *The Sources of Military Change: Culture, Politics, Technology, Making Sense of Global Security* (Boulder, CO: Lynne Rienner Publishers, 2002); Colin Elman, "The Logic of Emulation: The Diffusion of Military Practices in the International System" (Ph.D. Dissertation, Columbia University, 1999); Joao Resende-Santos, "Anarchy and the Emulation of Military Systems: Military Organization and Technology in South America, 1870-1930," *Security Studies* 5, No. 3 (1996):193–260.; Thomas Z. Zarzecki, *Arms Diffusion: The Spread of Military Innovations in the International System* (London, UK: Routledge, 2002).; Emily O. Goldman, "Cultural Foundations of Military Diffusion," *Review of International Studies* 32, No. 1. (2006): 69–91.; Paul Bracken, "Non-Standard Models of the Diffusion of Military Technologies," *Defense Analysis* 14, No. 2 (1998): 101–114.

On adaptation, see Meir Finkel, *On Flexibility: Recovery from Technological and Doctrine Surprise on the Battlefield* (Stanford, CA: Stanford University Press, 2011).; Adam M. Jungdahl and Julia M. Macdonald, "Innovation Inhibitors in War: Overcoming Obstacles in the Pursuit of Military Effectiveness," *Journal of Strategic Studies* 38, No. 4 (2015): 467-499.; Nina Kollars. "Military Innovation's Dialectic: Gun Trucks and Rapid Acquisition." *Security Studies* 23, No. 4 (2014): 787–813.; Frank G. Hoffman, *Mars Adapting: Military Change During War* (Annapolis, MD: Naval Institute Press, 2021).; Raphael Marcus, "Military Innovation and Tactical Adaptation in the Israel–Hezbollah Conflict: The Institutionalization of Lesson-Learning in the IDF," *Journal of Strategic Studies* 38, No. 4 (2015): 344-370.; Michael A. Hunzeker, *Dying to Learn: Wartime Lessons from the Western Front* (Ithaca, NY: Cornell University Press, 2021).

why the subset of new technologies that it typically focuses on were developed, or how effective the new technologies, with or without accompanying new doctrine, were independent of later wartime experiences that can be determined on the basis of hindsight. Limited though the military innovations literature in multiple respects, including its diffusionist tendencies, which results in frequent oversight as to the potential for countermeasures as opposed to myopic emulation, it nevertheless constitutes one of the primary areas of international relations scholarship to examine the implications of advances in military technology.

Although much of international relations scholarship does not concern itself with military technology or technological change in the same, more direct manner as the literatures on offense-defense theory and military innovations, many international relations scholars nevertheless examine the implications of military power in some way, shape, or form and are as such concerned with shifts in the military power of states. Military power, broadly defined, is, after all, one of the primary independent variables in scholarship on the causes of war, patterns of alliance and alignment, arms racing, and crisis dynamics, among other foci of research. Some of the most influential concepts in international relations scholarship, such as the security dilemma and polarity, are respectively predicated upon a shift in military power in a state dyad or are a characterization of the distribution of military power in the international system and shifts therein. Such is the centrality of military power and shifts in military power in international relations scholarship that writings on topics as far afield as voting patterns in international organizations and patterns of trade test the possibility that the distribution of military power can account for variation in these phenomena of interest.

While international relations scholars have made considerable efforts toward understanding the implications, broadly conceptualized, of shifts in the military power of states including shifts in which technological change plays a part, many writings have a major shortcoming in conceptualizing military power and shifts in military power in a highly abstracted manner and this affects how much of international relations scholarship views the *military effects* – as opposed to political and/or diplomatic effects – of technological change for international politics. By and large, international relations scholarship focuses on the effects of military power rather than its causes, manifestations, or interactions. As a result, the *military effects* of a shift in military power tend to be simply assumed with many writings focusing on the political and diplomatic implications of a presumed shift in military power. Given the relative dearth of scholarship on matters of tactics, logistics, operational planning, force structure, technology development, and, more generally, the conduct of warfare,[11] this is not surprising, but the consequence of

[11] This is increasingly changing primarily as a result of the growing literature on the sources of military effectiveness. See, for example, Caitlin Talmadge, *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes* (Ithaca, NY: Cornell University Press, 2015).; Ryan Grauer, *Commanding Military Power: Organizing for Victory and Defeat on the Battlefield* (Cambridge, UK: Cambridge University Press, 2016).; Michael A. Hunzeker, *Dying to Learn: Wartime Lessons from the Western Front* (Ithaca, NY: Cornell University Press, 2021).; Jasen Castillo, *Endurance and War: The National Sources of Military Cohesion* (Stanford, CA: Stanford University Press, 2014).; Rosella Cappella Zielinski and Ryan Grauer, "Organizing for Performance: Coalition Effectiveness on the Battlefield," *European Journal of International Relations* 26, No. 4 (2020): 953-978.; Stephen Biddle and Stephen Long. "Democracy and Military Effectiveness: A Deeper Look," *Journal of Conflict Resolution* 48, No. 4 (2004): 525–46.; Stephen Biddle and Robert Zirkle. "Technology, Civil-Military Relations, and Warfare in the Developing World," *Journal of Strategic Studies* 19, No. 2 (1996): 171– 212.; Michael Beckley, "Economic Development and Military Effectiveness," *Journal of Strategic Studies* 33, No. 1 (2010): 43-79.; Ryan Grauer and Michael Horowitz, "What Determines Military Victory? Testing the Modern System," *Security Studies* 21, No. 1 (2012): 83–112.; Allan R. Millett, Williamson Murray and Kenneth H. Watman, "The Effectiveness of Military Organizations," *International Security* 11, No. 1 (1986): 37-71.; Allan R. Millett and Williamson Murray, Editors., *Military Effectiveness*, Vols. 1–3 (Boston: Allen and Unwin, 1988); Risa Brooks, "An Autocracy at War: Explaining Egypt's Military Effectiveness, 1967 and 1973," *Security Studies* 15, No. 3 (2006): 396–430.; Dan Reiter, "Nationalism and Military Effectiveness: Post-Meiji Japan," in *Creating Military Power: The Sources of Military Effectiveness*, edited by Risa A. Brooks and Elizabeth A. Stanley (Stanford, CA: Stanford University Press, 2007); Barry R. Posen, "Nationalism, the Mass Army, and Military Power," *International Security* 18, No. 2 (Fall 1993): 80-124.; Dan Reiter and Allan Stam, "Democracy and Battlefield Military Effectiveness," *Journal of Conflict Resolution* 42, No. 3 (1998): 377–

this tendency toward abstraction is that much of international relations scholarship does not identify concrete mechanisms through which shifts in military power transpire. As a result, much of international relations scholarship is poorly positioned to assess the *military effects* – vice the political and diplomatic implications – of shifts in military power and the materiality of shifts that it identifies including those ostensibly driven by technological change.

Perhaps the greatest consequence of this tendency toward abstract conceptions and measures of military power is the frequent, though certainly not universal, tendency toward conflation of the military *potential* of states with the *fielded* military capabilities of states. While a state experiencing an increase in its military *potential* has an increased wherewithal to raise, train, and equip *additional* military capabilities (forces) and sustain those that already exist, it does not necessarily have additional *fielded* military capabilities at its immediate or near-term disposal. A state experiencing an increase in its *fielded* military capabilities, in contrast, has an increased capacity to undertake one or more (real-world) military missions either on its own accord (i.e., monadically – without an adversary) or in opposition to one or more adversaries (i.e., dyadically).

89.; David A. Lake, "Powerful Pacifists: Democratic States and War," American Political Science Review 80, No. 1 (1992): 24-37.; Stephen Biddle and Stephen Long, "Democracy and Military Effectiveness: A Closer Look," *Journal of Conflict Resolution* 48, No. 4 (2004): 525–46.; Michael C. Desch, *Power and Military Effectiveness: The Fallacy of Democratic Triumphalism* (Baltimore, MD: Johns Hopkins University Press, 2008).
Alexander B. Downes, "How Smart and Tough Are Democracies? Reassessing Theories of Democratic Victory in War," *International Security* 33, **N**o. 4 (2009)**:** 9-51.; Kenneth Pollack, *Arabs at War* (Lincoln, NE: University of Nebraska Press, 2002).; Stephen P. Rosen, *Societies and Military Power: India and Its Armies* (Ithaca, NY: Cornell University Press, 1996).; Stephen P. Rosen, "Military Effectiveness: Why Society Matters," *International Security* 19, No. 4 (1995): 5–31.; Kenneth M. Pollack, "The Influence of Culture on Arab Military Effectiveness," (Ph.D. dissertation, Massachusetts Institute of Technology, 1996); Risa A. Brooks and Elizabeth A. Stanley, Editors, *Creating Military Power: The Sources of Military Effectiveness* (Stanford, CA: Stanford University Press, 2007)

An increase in the military potential of a state may lead to an increase in its fielded military capabilities later in time, and this is not unimportant, but an increase in *fielded* military capabilities is not an intrinsic effect of an increase in military potential, and any effects tend not to be immediate. An increase in the gross domestic product (GDP), for example, or an increase in population or even an increase in military expenditures – three (proxy) measures of military power that are regularly invoked in international relations scholarship much as these are in policy debates – does not necessarily translate into an increase in *fielded* military capabilities of a state or affect the balance of fielded military capabilities between two states. For example, a state may or may not allocate its newfound economic capacity toward its fielded military capabilities and the effects on its fielded capabilities of a decision to do so will in any event be lagged. Stated differently, an increase in State A's military potential may facilitate a shift in its fielded military capabilities at a later date but may have no intrinsic and immediate implications for the State A-State B military balance. The widespread – but not universal – conflation of the military potential of states with their fielded military capabilities reflects upon the tendency of international relations scholarship to both treat military power as an abstraction and not identify the mechanisms through which shifts in fielded military capabilities are brought into being.

International relations scholars are, by and large, well aware of the limitations of existing – and readily accessible – concepts and measures of military power, and current trends in scholarship reflect a greater interest in the conduct of wars and the effects of pre-war and intra-war decisions concerning training, command, logistics, and so forth. For the most part, however, international relations scholarship continues to rely upon

abstract conceptions and measures of military power and continues to substitute military potential for fielded military capabilities when the task at hand requires the latter. The allure of readily available and parsimonious conceptions and measures of military power is understandable, but it is not without cost even as more abstract conceptions and measures are more suitable for some purposes than others. The practice is most problematic when scholars characterize an increase in one state's capacity to undertake military missions and more generally inflict harm onto another without identifying concrete indicators or specifying mechanisms. In such cases, scholars cannot simply conflate an increase in a state's military potential with an increase in its fielded military capabilities; an increase in State A's GDP, CINC score, or even its military expenditure, for example, does not necessarily increase State A's fielded military capabilities and does not therefore necessarily reduce State B's fielded military capabilities (or State C's, etc.) in relation to those of State A. If an observer claims that State B experiences a decrease in its fielded military capabilities relative to State A in such a scenario, or claims that State B has a material basis for an *immediate* increase in its perceived insecurity, then it is incumbent upon the observer to accurately characterize the *military effects* – as opposed to political and/or diplomatic implications – of this claimed shift in fielded capabilities, and this requires the accurate identification and explication of its causes and mechanisms. *Something* must be responsible for a claimed non-zero shift in *fielded* military capabilities of a state or else increased threat perceptions would lack a material basis.

Although *something* must be responsible for shifts in the fielded military capabilities of states, it is important to emphasize that this does not necessarily mean that all shifts in threat perceptions must have a material basis rooted in changes to fielded

military capabilities. One state may, after all, consider another state to be more threatening from one day to the next for any number of reasons *even if there are no substantive changes in the balance of fielded military capabilities*.[12] A change in political leadership, a political speech, new legislation, and so forth are examples of day-to-day politically salient events that can lead one state to revise or update its threat perceptions about another state even in the absence of any shift, whether real or perceived, in the balance of fielded military capabilities between said states (or any martial action). If, however, real-world policymakers or scholars engaged in theoretical research or policy analysis claim that State A finds State B more threatening *because of a shift (gain) in State B's military capabilities* relative to State A, then there must be a discernable and explicable substantive and material basis to such a characterization. Otherwise, it could simply be that State A considers State B more threatening on the basis of State B's *existing* military capabilities and this could simply be a manifestation of the increased salience of the *existing* State A-State B military balance and not a material – that is substantive – shift in this military balance. The fielded military capabilities of states are not mere abstractions and an inability to identify and explicate the material basis of change in threat perceptions driven by claimed shifts in military capabilities could make the revised threat perception indistinct from a hallucination.

In this dissertation project, I will investigate the relationship between technological change and shifts in the fielded military capabilities of states. Specifically, I will examine whether technological change, by which I specifically refer to advances in military technology, is a suitable candidate for the role of a prime mover engendering

[12] On non-material factors shaping threat perceptions, consider, for example, Jennifer Mitzen, "Ontological Security in World Politics: State Identity and the Security Dilemma," *European Journal of International Relations*, Vol. 12, No. 3 (2006): 341-370.

sizeable and abrupt shifts – gains – in the fielded military capabilities of states. I will first explain why international relations scholarship requires the explication of a prime mover capable of engendering shifts – gains – in the fielded military capabilities of states that are *both sizeable and abrupt*. I will then briefly consider candidates for the role of a prime mover identified in international relations scholarship and explain why most of these with the notable exception of technological change are unsuitable candidates in terms of causing shifts in fielded military capabilities that are both sizeable and abrupt. After developing these arguments in Chapter Two, I will then develop an analytical framework in Chapter Three that examines both monadic and dyadic mechanisms through which technological change can engender shifts in the fielded military capabilities of states. After employing this analytical framework to identify reasons as to why technological change is an unsuitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states, I will undertake empirical work to discern whether the historical record accords with said analytical framework. Through empirical and analytical work focusing on advances in submarine technology in the ca. 1800-1945 timeframe, the chapters of which constitute Part Two of this dissertation, and advances in uncrewed aircraft technology in the ca. 1910-2020 timeframe, the chapters of which constitute Part Three of this dissertation, I will present evidence that there is little in the historical record supporting the notion that technological change tends to engender sizeable and abrupt shifts in the fielded military capabilities of states. As a result of these preliminary findings, international relations scholarship remains in search of a prime mover well suited to engendering sizeable and

abrupt shifts in the fielded military capabilities of states that are required to get the proverbial ball moving in many influential areas of scholarship.

Before proceeding, it is important to explicate that this dissertation project exclusively examines the *military effects* of technological change which is to say the pace and magnitude of shifts in fielded military capabilities engendered by advances in military technology. This dissertation does not, in other words, make any claims as to the political and/or diplomatic implications, or other plausible sets of implications, of advances in military technology, which tends to be the primary focus of international relations scholarship on the matter. Logically, the political and diplomatic implications of advances in military technology should reflect upon the military effects of technological change, which is to say that the perceived political and diplomatic implications, as understood by international relations scholars and others, should be informed by an accurate appraisal of the pace and magnitude of shifts engendered in fielded military by technological change. This is, of course, how scholars and analysts should ideally examine international politics, whether contemporary, historical, or prospective. The historical record and more recent experiences, however, suggest the existence of many episodes in which a wide chasm between the real-world political and diplomatic implications of an advance in military technology – which tend to be shaped by the perceived and typically prospective effects of technological change on military capabilities if about technological change at all – and the observable military effects of advances in military technology/technological change – which are more concretely grounded in the pace and magnitude of shifts engendered in fielded military capabilities at a given point in time. In principle, the military effects of technological change are

straightforward to discern and can be understood in purely monadic terms but the real-world effects of technological change are best understood in dyadic terms given that the net effects of technological change on the fielded military capabilities of states in a state dyad are invariably shaped – coproduced – by both the actions taken and forgone by states in relation to one another as well as the highly variant military-geographical context in which a new technology is introduced and employed. Stated differently, much as shifts in threat perceptions do not necessarily require a material basis to affect international politics, perceptions of the political and/or diplomatic implications of technological change can take on a life of their own and diverge from the materially discernable military effects of technological change.[13]

Although this dissertation focuses on the *military effects* of technological change and does not therefore examine the political and diplomatic implications of (presumed) shifts in fielded military capabilities that much of international relations scholarship focuses on, this is not to suggest let alone claim the primacy of military considerations of what are profoundly political matters. There can, in other words, exist a significant gap between perceptions of shifts in a military balance and what transpires. As a result, even a technological advance that is not successfully developed and fielded can have political and/or diplomatic effects on real-world international politics but there is nevertheless an analytical case to be made for international relations scholars, among others, to make the distinction. Even if one takes the view that the political and diplomatic implications of technological change, which are shaped by perceptions of military balances independent of the materiality of such perceptions, are ultimately of foremost importance, scholars of

[13] Prominent examples of this dynamic can be found in the context of concerns about aerial bombardment in the United Kingdom during the Interwar Period and concerns about first a "bomber gap" and later a "missile" gap in the United States during the early Cold War.

international politics require a framework through which to assess the degree to which perceptions of shifts in fielded military capabilities accord with the substantive effects of technological change on the fielded military capabilities of states. This dissertation develops an analytical framework through which to discern the substantive effects of technological change on fielded military capabilities and assesses whether said analytical framework itself accords with the empirical record and thereby endeavors to establish a foundation that future research can build upon.

Chapter 2: Shifts in Fielded Military Capabilities and the Necessity of a Prime Mover

Although most international relations scholars irregularly focus on fielded military capabilities in any particular detail, a significant portion of scholarship requires *something* to drive shifts in the fielded military capabilities of states and is predicated upon the notion that such shifts in fielded military capabilities transpire and are of consequence for international politics. Without what can be characterized as a prime mover to drive – to be the driving force of – such shifts in the fielded military capabilities of states, variance – temporal and otherwise – in threat perceptions would lack a material basis.

To appreciate the importance of determining what drives shifts in the fielded military capabilities of states – situations in which one state benefits from a gain in fielded capabilities in relation to and at the expense of another, it is productive to turn to the influential concept of a security dilemma. While there are multiple formulations of a security dilemma,[14] at its core the concept refers to a two-actor model in which State A's actions to improve "its security," which is typically a euphemism for improving military capabilities, result in State B feeling threatened on the basis of a perception that State A's

[14] See John H. Herz, "Idealist Internationalism and the Security Dilemma," *World Politics* 2, No. 2 (1950): 157-180.; Herbert Butterfield, *History and Human Relations* (London, UK: Collins, 1951).; Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30, No. 2 (1978): 167–214.; Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976).; Charles L. Glaser, "The Security Dilemma Revisited," *World Politics* 50, No. 1 (1997): 171- 201.; David A. Lake, "Escape from the State of Nature: Authority and Hierarchy in World Politics." *International Security* 32, No. 1 (2007): 47-79; Charles L. Glaser, *Rational Theory of International Politics* (Princeton, NJ: Princeton University Press, 2011).; Charles L. Glaser, "The Security Dilemma Revisited," *World Politics* 50, No. 1 (1997): 171- 201.; Randall L. Schweller, "Neorealism's Status-Quo Bias: What Security Dilemma?" *Security Studies* 5, No. 3 (1996): 90-121.; Shiping Tang, "The Security Dilemma: A Conceptual Analysis," *Security Studies* 18, No. 3 (2009): 587-623.; Barry R. Posen, "The Security Dilemma and Ethnic Conflict," *Survival* 35, No. 1 (1993): 27-47.

measures undermine its own (i.e., B's) "security." As a result, State B may undertake efforts to improve "its security" in response and this may in turn result in State A experiencing the same adjustment to its threat perceptions such that it undertakes further measures to improve "its security" in turn. Figure 1 lays out the core mechanisms of a security dilemma beginning with a shift in a state's fielded military capabilities.

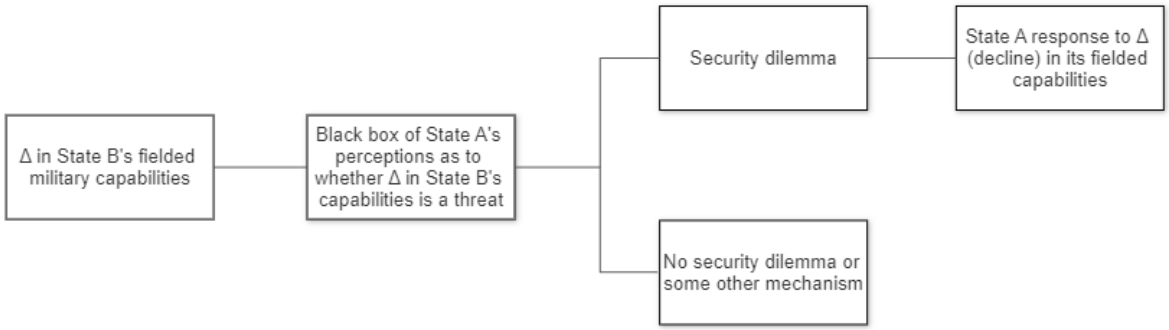

Figure 1. Shifts in Fielded Military Capabilities and Security Dilemma Dynamics

Over the decades, international relations scholars have both built upon the core logic of a security dilemma and criticized it in ways that, by offering additional mechanisms, suggest means through which security dilemmas can be ameliorated if not entirely avoided. Robert Jervis, for example, influentially argued that two factors – whether the offense or defense has an advantage and whether offensive/defensive

postures are distinguishable – affect the severity of a security dilemma.[15] More recently, Charles Glaser has argued that the severity of a security dilemma is determined by the interaction of material variables with information variables.[16] Critics of the security dilemma framework similarly propose additional mechanisms of their own. Some scholars writing in the liberal tradition of international relations theory, for example, argue that regime type affects threat perceptions such that two democracies may not find one another as threatening whether as a result of democratic norms, institutional constraints, or some other mechanism.[17] Similarly, some scholars writing in the constructivist tradition of international relations theory highlight the role of identities in shaping threat perceptions.[18] If threats, including those animating a security dilemma dynamic, have a purely material basis, the argument goes, then the United States should be as threatened by British and French nuclear weapons as it is by the similarly sized (thus far) Chinese nuclear force. That it is not, constructivist critics contend, suggests that there are extra-material factors at work (which need not necessarily be those advanced by constructivist scholars) and that the (core) security dilemma framework is therefore underspecified. Figure 2 highlights how both proponents of the security dilemma framework and their critics have tended to add to the core framework rather than

[15] Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30, No. 2 (1978): 167–214.
[16] See, for example, Charles L. Glaser, "Realists as Optimists: Cooperation as Self Help," *International Security* 19, No. 3 (1994/1995): 50-90.; Charles L. Glaser, "The Security Dilemma Revisited," *World Politics* 50, No. 1 (1997): 171- 201.; Charles L. Glaser, *Rational Theory of International Politics* (Princeton, NJ: Princeton University Press, 2011).
[17] Seth Weinberger, "Institutional Signaling and the Origin of the Cold War," *Security Studies* 12, No. 4 (2004): 80-115.; Michael W. Doyle, "Kant, Liberal Legacies, and Foreign Affairs," *Philosophy and Public Affairs* 12, No. 3 (1983): 205-235.; Michael W. Doyle, "Kant, Liberal Legacies, and Foreign Affairs, Part 2," *Philosophy and Public Affairs* 12, No. 3 (1983): 323-353.
[18] See, for example, Jennifer Mitzen, "Ontological Security in World Politics: State Identity and the Security Dilemma," *European Journal of International Relations*, Vol. 12, No. 3 (2006): 341-370.; Peter J. Katzenstein, Editor. *The Culture of National Security: Norms and Identity in World Politics* (New York: Columbia University Press, 1996); Alexander Wendt, "Anarchy Is What States Make of It," *International Organization* 46, No. 2 (1992): 391-425.

questioning the starting assumptions or adding mechanisms situated before the black box of state perception (on the causal chain) that is the central focus of debate.

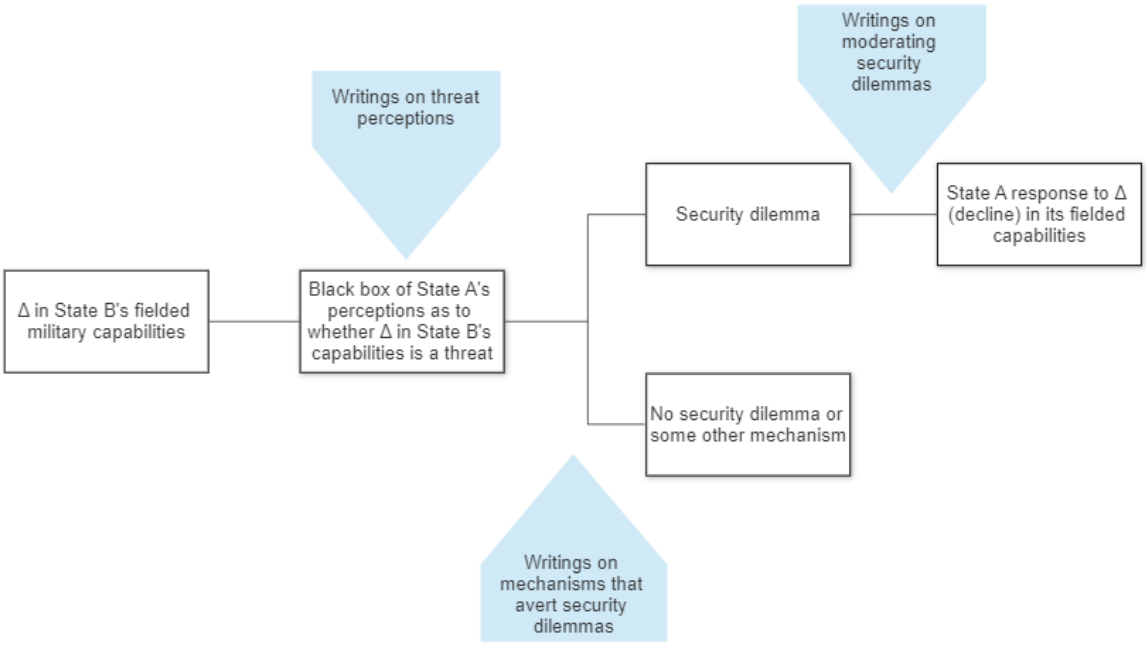

Figure 2. Gaps in International Relations Scholarship

What is striking about this influential area of international relations scholarship is that both proponents of some formulation of a security dilemma framework and the critics thereof tend to start with the same assumption – that one state is the beneficiary of a (presumed) shift in its military capabilities and that another country experiences a decline in its military capabilities – and essentially take it for granted that the (presumed) capability shift is in *purely military terms* – as opposed to political and/or diplomatic terms – significant enough for the other state to have substantive material grounds for a heightened sense of insecurity. Stated differently, both proponents of the security dilemma framework and its critics take the presumed initial capability shift for granted

while typically not specifying how exactly the fielded military capabilities of a state change – increase or decrease – in either monadic or dyadic terms.

2.1: How Can States Increase Their Military Capabilities?

While the security dilemma framework, like much of international relations scholarship, begins with the starting assumption of a shift in the military capabilities of a state, it does not explain how this shift in capabilities is brought into being and this makes it difficult to discern the substantive and material implications of a gain, whether real or hypothetical, in State A's fielded military capabilities for State B. How can states obtain an increase in their fielded military capabilities?

The list of distinct actions that can plausibly increase a state's fielded military capabilities is long, but logically these efforts can be divided into two non-exclusive approaches. First, a state can allocate additional resources to its military. A state can, for example, procure and field additional military equipment (which may be manifestations of existing and/or new military technology), it can develop and introduce new military technology, increase the size of its military, and allocate additional resources toward training, readiness, and maintenance, and so forth. Alternatively, a state can undertake reforms to derive an increase in fielded military capabilities from its existing manpower and materiel – as well as forthcoming manpower and materiel. A state can, for example, adjust its military doctrine and/or military strategy, reorganize its existing military units, and retrain its existing military personnel. These two approaches – *resource injection* and *reform* – are not mutually exclusive and can be further disaggregated but nevertheless illustrate some of the different concrete approaches through which states can increase

their fielded military capabilities. Notably, these approaches do not include the standard array of concepts and associated measures of military power identified in much of international relations scholarship such as an increase in GDP, population size, or military expenditures because the latter affect the military *potential* of states and do not directly affect their fielded capabilities. To change its fielded military capabilities, a state will have to convert its newfound military potential toward fielded capabilities through resource injection and/or military reforms. Otherwise, there is no material shift – no substantive shift – in the balance of *fielded* military capabilities between the two states. Figure 3 adds these pathways to the causal chain of the core elements of a security dilemma framework.

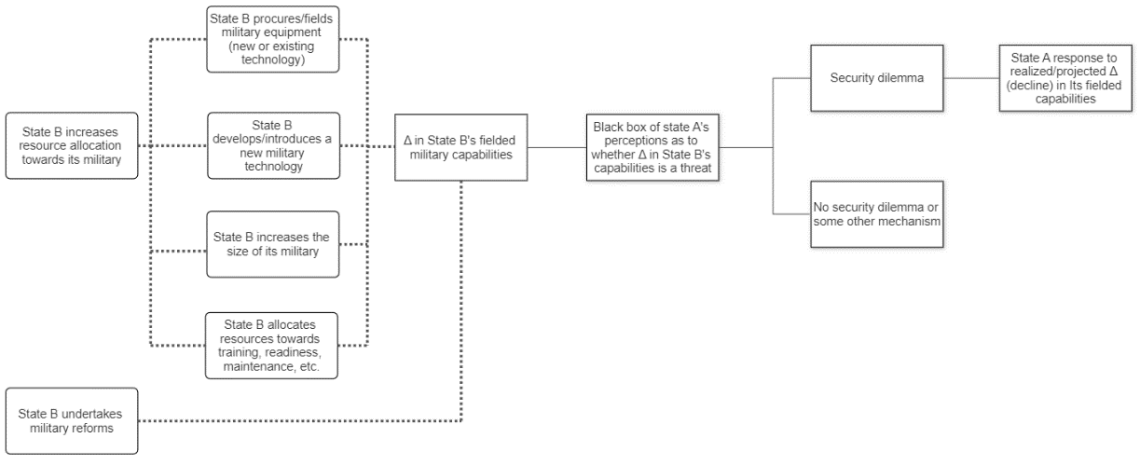

Figure 3. How States Expand Their Fielded Military Capabilities

Expanding the causal chain backward to account for *how* states can obtain an increase in their fielded capabilities allows us to develop a more concrete understanding of the mechanisms and the potential military effects of shifts in fielded military capabilities for international politics. Since a shift in fielded capabilities is not an abstract matter in the real world, this basic framework helps us to consider how other states may attempt to mitigate the potential effects of a decline in their military capabilities and this is logically prior to their threat perceptions. Stated differently, threat perceptions should logically take into account the capacity of a state to mitigate against the effects of what would otherwise be a decline in their fielded military capabilities, or else there would logically be little if any variance in threat perceptions. An increase in the size of an adversary state's military, for example, may lead states to pursue a different response than an increase in an adversary's military readiness. Similarly, doctrinal reforms undertaken by an adversary may drive a different response than an adversary's development of a new military technology. If there is variance in the degree to which the effects of a capability shift can be mitigated, then the military effects of a capability shift for international politics are mutable by the actions of human agents and are not therefore an intrinsic effect of a capability shift.

Although raising the question of how states can obtain an increase in their fielded military capabilities is a necessary first step, it brings us no closer to understanding the military effects of a shift in fielded military capabilities for international politics. Even if State A's undertakings – resource injection and/or reform – result in a non-zero shift in the State A-State B military balance, such a capability shift does not have intrinsically deleterious military effects for State B and, as such, has no automatic political and/or

diplomatic implications for international politics. To discern the implications of a non-zero shift in fielded military capabilities, we require an analytical framework that identifies the key parameters determining the significance of a capability shift for both the state pursuing the capability gains and other states that stand to experience a potential relative decline in their fielded military capabilities as a result. To this end, I propose that the implications of a non-zero shift in a state's fielded military capabilities logically vary in significance across at least two parameters: the *pace of change* in fielded capabilities (i.e., the timeframe over which a shift in fielded capabilities is realized) and the *overall magnitude of change* in fielded capabilities (i.e., the total displacement of a capability shift in nominal or absolute terms). The pace of change in fielded military capabilities can be abrupt (i.e., transpire in a shorter timeframe) or gradual (i.e., transpire over a longer timeframe) and the magnitude of change can be small or large. Given that each shift in fielded military capabilities can be characterized through these two parameters, there are four possible characterizations of the military effects of a non-zero shift in the fielded capabilities of states, and these are specified in Figure 4.

		Pace of Change in Fielded Military Capabilities	
		Gradual	**Abrupt**
Magnitude of Change in Fielded Military Capabilities	**Small**	Least effects	Moderated effects
	Large	Moderated effects	Greatest effects

Figure 4. The Interaction of the Pace and Magnitude of Changes in Fielded Military Capabilities

The four possible characterizations of the effects of shifts in fielded military capabilities based on the two parameters, the pace of change and the magnitude of change, are as follows: shifts in capabilities that are gradual and small, shifts that are abrupt and small, shifts that are gradual and large, and shifts in capabilities that are abrupt and large. These four characterizations reflect upon three distinct sets of military effects of non-zero shifts in a balance of fielded military capabilities. Shifts in fielded military capabilities that are small in magnitude, and which gradually unfold logically have the lowest possible military significance for international politics (the "Least effects" cell in Figure 4). These are cases in which other states not only encounter a modest (even if non-zero) decline in their military standing relative to another state but also have the most

time – the greatest potential opportunity – to undertake attempts to blunt if not negate the capability gains accrued to the other state (mechanisms through which this can be attempted will be examined later in this chapter).[19]

Cases in which (A) shifts in fielded capabilities are large in magnitude but gradually unfold and those in which (B) shifts in capabilities are small in magnitude but transpire at an abrupt pace are cases in which the military effects of a capability shift are moderated. In such cases, states may encounter a potentially significant decline in their fielded military capabilities vis-à-vis another state, but these shifts are nevertheless moderated at a given time for want of abruptness or magnitude. Because the abrupt shifts in fielded capabilities are small in magnitude and the large shifts in fielded capabilities gradually unfold in such cases, states are afforded greater opportunities to undertake measures through which to blunt if not negate the capability gains accrued to other states. Although the risk and time pressures that states encounter are inherently greater in the "moderated effects" cells than the "least effects" cell in Figure 4, the "moderated effects" cells do not reflect contexts in which advances in military technology engender capability shifts of unambiguously major significance to international politics insofar as there must be a material basis to threat perceptions rooted in shifting military capabilities.

Shifts in fielded military capabilities, whatever their cause, have the greatest possible implications for international politics when the capability shifts are both large and abruptly transpire – the "greatest effects" cell in Figure 4. These are cases in which

[19] Although shifts in capabilities that are small but abruptly transpire have the potential to become very significant/aggregate *over time*, this is most likely when the capability gains are characterized by compounding growth or exponential growth rates. If such conditions do not hold and if capability gains over time are characterized by multiple s-shaped curves reflective of diminishing gains absent a new major technological breakthrough (i.e., a dynamic of technological plateauing), then the aggregated effects over time take on a very different form.

military competition between states is best characterized as intense and wherein there is constant jostling for advantage to mitigate against the potential for a sizeable and abrupt relative decline in fielded military capabilities. Because the shifts in fielded capabilities abruptly unfold in such cases, states have less time to undertake efforts to blunt or negate the effects of technological change on their fielded military capabilities, and the risks associated with falling behind increase. Given that the magnitude of the shifts in capabilities are large/sizeable in such cases, states in a situation of the "greatest effects" cell are also pressured to look for advantages wherever these can be found, thereby instilling additional incentives for all states to continue the competitive dynamic in pursuit of capability gains in order to avoid a potentially significant decline in fielded military capabilities (i.e., a red queen hypothesis like dynamic in which states must constantly compete – run – to simply maintain their position and not fall behind).[20]

The framework of effects represented in Figure 4 is predicated upon the notion that both the pace and magnitude of shifts in fielded military capabilities are consequential. To illustrate why this framework and these two parameters are analytically productive, I will discuss the implications of notional shifts in fielded military capabilities resulting from technological change.

The condition of abruptness is a productive parameter for discerning the military significance of capability shifts because states have the potential to respond to other states' efforts to increase their fielded military capabilities such as when another state develops and fields a new military technology. If, for example, State A anticipates that

[20] Leigh Van Valen, "A New Evolutionary Law," *Evolutionary Theory* 1, No. 1 (1973): 1-30. While applied to evolutionary biology, the term red queen refers to Lewis Carroll's 1872 novel *Through the Looking-Glass, and What Alice Found There*.

State B will experience growth in its fielded military capabilities as a result of State B's pursuit of new technology, *technology x*, then State A has a variety of options to stave off a relative decline in its own fielded military capabilities. State A can itself *adopt* technology *x* (it may have already been doing so – an act of convergence – or would otherwise emulate State B).[21] State A can pursue some combination of technological and/or tactical-organizational-doctrinal *countermeasures* to technology *x*. State A can also pursue technologies *y* and *z* and so forth to obtain advantages elsewhere in the military balance and thereby attempt to *offset* the non-zero gains in fielded capabilities that State B may accrue from introducing technology *x*. State A – and State B as well as other states – can of course simultaneously pursue more than one of these options for these are not mutually exclusive. Moreover, the *interaction* of these competing and often varied efforts can create a steady flow of shifting incentive structures concerning the allocation of resources such that there may be multiple dynamic pathways for State A to productively respond to State B's introduction of technology *x* over time so as to blunt if not negate the effects of technological change on its fielded military capabilities relative to State B (the incentives may vary for other states, whether State C relative to State A, State A relative to State C, State C relative to State D, etc. and can also vary across time, space, and military-geographical contexts). Table 1 characterizes these non-exclusive approaches that a state can undertake in response to technological advances pursued by other states in order to moderate the net effects, if any, on its fielded military capabilities. These approaches encompass both the injection of additional resources toward military

[21] This distinction is important but often overlooked through the use of the term "emulation." Logically, one state emulates another only when both states were not independently pursuing a technology. Consequently, a distinction should be made between emulation and convergence.

capabilities and reforms aimed at obtaining increased capabilities with existing manpower and materiel.

Table 1. Adopt, Counter, and/or Offset: How States Can Respond to Technological Advances Pursued by Other States	
Adopt the (same) new technology	*Convergence* if both states are independently pursuing the new technology
	Emulation if one state pursues the new technology in response to another state first doing so
Counter the new technology	*Technological countermeasures* to blunt if not negate the effects of the new technology
	Organizational-doctrinal adaptations (countermeasures) to blunt if not negate the effects of the new technology
Offset the new technology	Attempt to supplant the new technology to offset the capability gains accrued to the other state
	Look for capability gains elsewhere in the military balance to offset the capability gains accrued to the other state through its pursuit of the new technology

Through these competitive dynamics – adopting, countering, offsetting – states have opportunities to blunt if not negate the gains in *fielded* military capabilities that other states seek through the development and deployment of new military technologies (or, more generally, some action aimed at increasing a state's fielded military capabilities). For such blunting or negation of capabilities to happen, however, states need time to respond. If State A's response to the new technology *x* begins only after State B starts to introduce – field – the new technology, then State A may experience a relative decline (i.e., a non-zero shift) in its fielded military capabilities vis-à-vis State B

even if, at a later date, it succeeds in blunting if not negating the implications of the new technology *x* on the State A-State B military balance. Consequently, the pace at which technological advances engender shifts in fielded military capabilities delineates the timeframe that states have to undertake efforts to avert a significant relative decline in a military balance before it transpires. All things equal, if shifts in fielded military capabilities engendered by technological change accrue gradually, then states have more time to respond, and the technological advance in question may have more limited effects on the military balance and, as such, less acute military significance for international politics. If, however, the gains in fielded military capabilities engendered by technological change accrue rapidly to the state introducing the new technology, then the technological change may have abrupt effects on the military balance and more acute military significance for international politics.

Given that State A can pursue a variety of approaches (i.e., adopt, counter, offset) in response to State B's pursuit of a notional new technology that has the potential of engendering shifts in the military balance, it is important to appreciate that shifts in the fielded military capabilities of states are not equal in terms of magnitude. The *military effects* of capability shifts, and as such the military implications for international politics, vary with the *interaction* of the pace of change (i.e., the timeframe over which a capability shift unfolds) with the magnitude of change even if the ultimate delta in engendered fielded capabilities is non-zero. While State A is likely to be better positioned – all things equal – to stave off a decline in its fielded military capabilities vis-à-vis State B if the latter gradually accrues the capability gains engendered by its fielding of new military technology, the implications for the State A-State B military balance will

necessarily vary with the gross or absolute magnitude of the shift in capabilities resulting from State B's introduction of the new military technology. If the shifts – gains – in State B's capabilities (relative to State A) transpire gradually in small increments over a long(er) timeframe, then State A has not only more time to respond but the implications of the (relative) decline in fielded military capabilities that it experiences relative to State B will be moderated in that these will transpire in smaller increments per unit of time. Depending on the specifics of the notional technological advancements in question and how State B deploys it within its capability set in relation to State A's capability set and other actions taken and forgone by each state in both monadic and dyadic terms, State A may ultimately experience a modest or perhaps even negligible decline in its fielded military capabilities relative to State B irrespective of the absolute capability gains that the new technology nominally offered State B. As such, even if the notional technological advance has overall *non-zero effects* on the State A-State B military balance, the practical military effects could be marginal at best and of limited military significance for international politics. Stated differently, we have no reason to assume that an abstracted non-zero shift in fielded military capabilities is intrinsically militarily significant or consequential even if it has major political and/or diplomatic implications.

Having illustrated this simplified framework for assessing the military effects of shifts in fielded military capabilities for international politics across two parameters (i.e., the pace and magnitude of change), we can return to the core framework of a security dilemma to appreciate the importance of identifying a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. The core framework of a security dilemma is, as previously explained, predicated upon the existence of *some shift*,

often unspecified, in one state's military power and concerns itself with the political and/or diplomatic implications thereof. Because international relations scholarship tends to abstractly conceptualize military power, it tends not to specify the nature of this starting capability shift which is to say its purely military effects or significance. In Figure 5, I add the approaches through which states can increase their military capabilities to the left of the security dilemma framework (without engaging in debates about the workings or political and/or diplomatic implications of security dilemmas). As explained over the preceding pages, the *military* effects of a non-zero shift in fielded military capabilities are coproduced by the actions taken and forgone by each state and are not therefore automatic, and Figure 4 reflects upon the prior postulation that there are four possible characterizations of effects when examined through the parameters of the pace of change and the magnitude of change in fielded military capabilities. Figure 5 inserts this framework for discerning the *military* effects between the black box of perceptions that writings on security dilemmas work to explain and the prior characterization of the concrete approaches that states can undertake to increase their fielded military capabilities

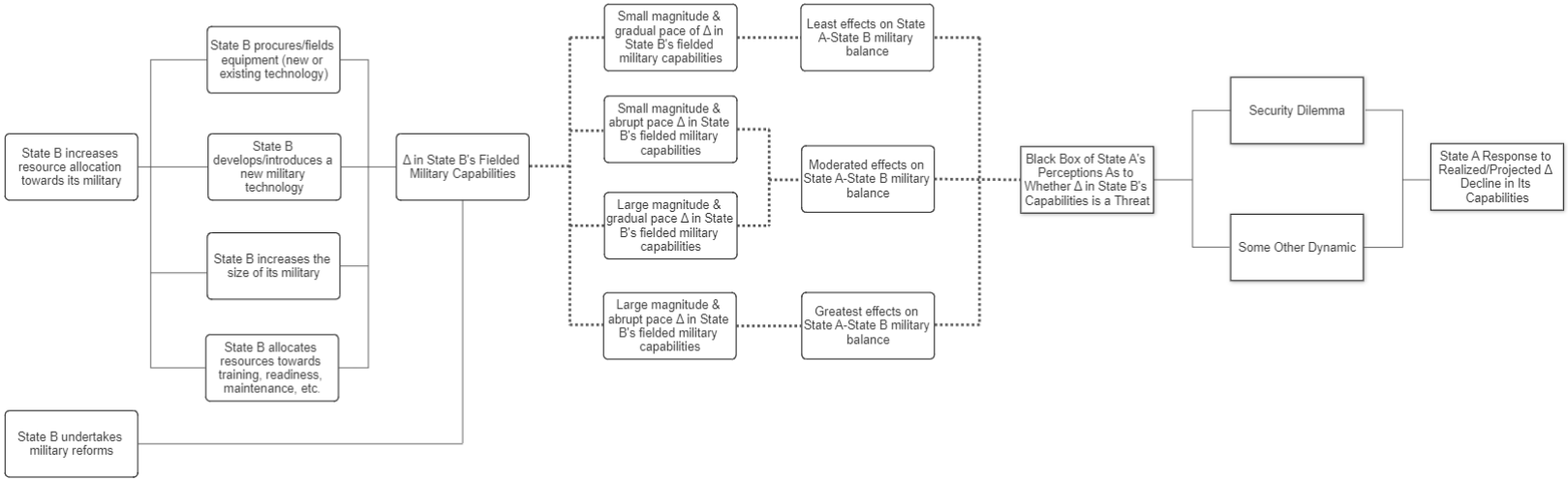

Figure 1. Fitting Adopting, Countering, and/or Offsetting Into the Security Dilemma Framework

Figure 5 highlights how shifts in fielded capabilities have no intrinsic military effects on international politics such that it is analytically unproductive to abstractly conceptualize shifts in military power. If nothing else, the *military* effects of shifts in fielded capabilities vary with the interplay of the pace and magnitude of change. Although states invariably form threat perceptions on the basis of multiple considerations of which the assessed military effects of a perceived shift in another state's fielded military capabilities are but one – and notwithstanding the possibility that the assessed military effects can be subordinate to various political and/or diplomatic considerations, it is nevertheless important for international relations scholars to make a distinction between the black box of threat perceptions and the substantive, that is material, extent of shifts in fielded capabilities. Even if a state experiences heightened threat perceptions, international relations scholars need to identify indicators of a substantive and material shift in the fielded military capabilities of one state in relation to another.

The pace and magnitude of shifts in capabilities are one example of such substantive indicators that international relations scholarship can benefit from. If, for example, State A declares its intentions to field one or more aircraft carriers, this results in *zero* effects on the State A-State B military balance until the carriers are developed, built, and deployed (independent of human perceptions), which is to say that there are likely to be *zero* effects on the dyadic military balance for several if not many years. Similarly, if State A announces an increase in military spending, this results in zero effects on the State A-State B military balance until the additional resources are allocated toward, and ultimately converted into, an expansion in fielded military capabilities. If this involves developing and deploying a new military technology, then the effects on the

balance of fielded military capabilities are zero until the new technology is developed, built, and deployed. Once said new technology is deployed, the *net* effects of any shift in capabilities are coproduced with the actions taken and forgone by an adversary and the degree to which it succeeds in any efforts to adopt, counter, and offset the new technology to blunt if not negate its effects on the dyadic military balance. This is not, of course, the only way to conceptualize the substantive implications of a shift in a military balance. State A may, for example, undertake an action that offers a gain in military capabilities, such as deploying a new military technology, but this could be a wasting asset in the sense that the newly deployed technology is already being displaced by technological change, which has invariably been the ultimate outcome across human history. In such a situation, the effects on a military balance may be non-zero but could nevertheless have a short shelf life and more generally serve as a wasting asset. Similarly, a state may introduce a new technology that is highly susceptible to countermeasures such that the effects on the military balance are non-zero but are readily reversible and therefore ephemeral.

What these examples highlight is that there is much to consider about the military effects of a shift in military capabilities independent of the discussions of the political and diplomatic implications that international relations scholars primarily concern themselves with. Unless the shifts engendered in fielded military capabilities by technological advances (or any other cause) are *both* sizeable and abrupt, the effects of a shift in capabilities may be moderated for both monadic reasons – for want of pace and/or magnitude of effects – and dyadic reasons – because other states can adopt, counter, and/or offset the capability gains accrued by a technological advance (or both given that

these are not mutually exclusive). For these reasons, international relations scholarship does not simply require something amounting to a prime mover to drive non-zero shifts in the fielded military capabilities of states; it requires a prime mover that engenders shifts in capabilities that are *both sizeable and abrupt*. It is only in such cases that states have little time to respond and therefore encounter a heightened risk that their efforts toward blunting if not countering a capability shift will not succeed in averting a decline in fielded military capabilities.

2.1 Candidates for a Prime Mover

Over the previous pages, I addressed the issue of how states can increase their fielded capabilities and offered two parameters through which international relations scholars can discern the *military effects* of a shift in fielded capabilities. Given that *something* must bring about a shift in fielded capabilities, what is the underlying cause? What candidates does international relations scholarship offer for the role of a prime mover affecting the fielded capabilities of states and are these candidates suited to the role of engendering shifts in fielded military capabilities that can be characterized as both sizeable and abrupt?

Although international relations scholars have developed numerous conceptualizations and metrics of military power including those that more broadly focus on "national power,"[1] consensus has often proven elusive and there is widespread recognition that some of the metrics, and as such the concepts that informed the creation

[1] See, for example, John J. Mearsheimer, *The Tragedy of Great Power Politics* (New York: Norton, 2001).; Michael Beckley, "The Power of Nations: Measuring What Matters," *International Security* 43, No. 2 (2018): 7-44.

of these metrics, leave much to be desired. Many writings focusing on inter-state security dynamics, including the influential literature on security dilemmas and interrelated literatures on arms racing, polarity, and the causes of war, place a heavy emphasis on military power and shifts in military power. To conceptualize and measure military power, scholars engaging in these literatures frequently refer to military expenditures, force structure (i.e., equipment counts and/or personnel counts), developments in military technology, and, less frequently, developments in the areas of tactics, force employment, and doctrine (this latter grouping is typically, but not exclusively, examined in the context of technological change). Although there is no consensus as to how the military power of states should be conceptualized and measured,[2] scholars contributing to this literature tend to appreciate that military power is itself a product of other sources of "national power" and is as such rooted in the economic capacities of states and demography, among other things. As such, many writings, including those referring to the aforementioned conceptualizations and metrics of military power, tend to also refer to demographic data and economic measures such as gross domestic product which, as previously explained, determine a state's military potential but are not necessarily reflective of its fielded military capabilities. To address the challenge of aggregating multiple disparate concepts and measures, scholars frequently use composite metrics such as the CINC score as a proxy measure for the military power of states. While these conceptualizations and approaches to measuring military power focus on individual states and are used to compare states, scholars writing on inter-state security dynamics tend to

[2] Beckley, "The Power of Nations."; Stephen G. Brooks and William C. Wohlforth, "The Rise and Fall of the Great Powers in the Twenty-First Century: China's Rise and the Fate of America's Global Position," *International Security* 40, No. 3 (Winter 2015-2016): 7-53.

place a heavy emphasis on the possibility that states can form alliances to pool their military power against one or more other states.

Over the decades, international relations scholars have identified a large number of potential causes of military power and factors that contribute to temporal and/or cross-station variation in military power. Scholars have examined the effects of, among other things, demographic change, economic change, diplomatic change, political change, (non-technological) military change, and technological change on the military potential and/or fielded military capabilities of states. These constitute the primary candidates for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. This dissertation project is neither intended to explain the threat perceptions of states nor is it intended to assess the mechanisms, or the relative importance of the causes or factors identified in international relations scholarship. The present objective is far more modest: to briefly summarize the main causes or factors identified in existing writings and examine whether these are suitable candidates for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. With a working list of candidates for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities in hand, the remainder of this chapter will examine these candidates for prime mover in order to address two questions. First, is there a direct mechanism through which change in a candidate area engenders shifts in the fielded military capabilities – vice military potential – of a state? Second, does change in this candidate area lend itself to driving shifts in fielded military capabilities that are both abrupt and sizeable?

To be a suitable candidate for the role of a prime mover, there must be a direct mechanism for translating changes in a given area into shifts in fielded military capabilities. Without a direct mechanism, the relationship between change in a given area and fielded military capabilities would be unclear, and there would be little reason for international relations scholars to confidently refer to a change in a given area when discussing a shift in the *fielded* military capabilities of states. A direct mechanism is not, however, wholly adequate (i.e., it is necessary but not sufficient), because international relations scholarship requires *something* that engenders significant shifts in capabilities that are both sizeable and abrupt. As explained earlier, if the shifts in fielded military capabilities are not both sizeable and abrupt, then – all things equal – states have more time to attempt to blunt if not negate the effects of shifts in fielded military capabilities. If this happens, then the *net* effects of the shift in fielded military capabilities can be moderated to such a degree as to be of limited significance for international politics.

2.2.1 Demographic Change

Although it is infrequently the focus of international relations scholarship, demography looms large in discussions of both the military potential and fielded military capabilities of states.[3] In principle, demographic change can engender sizeable shifts in the fielded military capabilities of states. Sustained population growth, for example, can increase the pool of manpower available to a state for military service and can simultaneously be allocated to the labor force so as to both fund military capabilities and, through the

[3] For a notable counterexample in which demography is directly examined, consider Mark L. Haas, "A Geriatric Peace? The Future of U.S. Power in a World of Aging Populations," International Security 32, No. 1 (2007): 112-147.

defense industry and other segments of the economy, equip and sustain military capabilities. Important though demographic change ultimately is – with time – it nevertheless constitutes the textbook case of lagged effects in that even a baby boom does not translate into any appreciable increase in manpower available for military purposes, let alone trained, equipped, and deployed manpower, for at least 15-20 years. As a result, demographic change does not lend itself to engendering shifts in fielded military capabilities that are both sizeable *and* abrupt. Moreover, demographic change lacks a direct mechanism through which to increase fielded military capabilities. Whatever the rate of population growth and the aggregate gains in a state's population, untrained manpower has been of decreasing importance over recent centuries in terms of its impact on both the military potential and fielded military capabilities of states.[4] Without adequate education, training, and industrial capacity, among other things, untrained manpower can only be employed in an ever-decreasing subset of combat roles and can be extremely ineffective when employed against more proficient and/or better-equipped adversaries. For these reasons, it is almost inconceivable that a security dilemma like dynamic, which begins with a shift (gain) in one state's military capabilities relative to another, could be initiated by demographic change serving in the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

2.2.2 Economic Change

While demographic change and population size delineate the maximum potential size of a state's labor force and the maximum potential size of its military (in terms of personnel),

[4] This point is increasingly emphasized in the military effectiveness literature.

personnel must be fed, equipped, undergo training, and so forth. The maximum practical size of a state's military is in practice, therefore, determined not by a state's demography but by its economy and economic capacity more generally. As a result, economic changes that increase the resources that a state can make available to its military can, in principle, affect the fielded military capabilities of states.

Economic changes are undoubtedly important for international politics but nevertheless lack a direct mechanism for increasing the fielded military capabilities of states. An increase in economic output, for example, has no intrinsic bearing on a military balance even as it increases the military potential of a state and thereby may provide conditions for shifts in its fielded military capabilities at a later date. The effects of economic change also tend to unfold gradually, and this is particularly the case with economic change driven by (non-military) technological development, institutional reform, and the more general advancement of human knowledge. Many major technological advances identified in economic history, for example, typically undergo many years of gradual development before these became suitable for most use cases and therefore resulted in heavily lagged even if ultimately significant (non-zero) implications for the economy and, as such, for international politics. Gains in productivity through the combination of innovation and accumulating experience, as well as institutional changes, similarly tend not to emerge overnight and therefore have greater implications over the long run.[5]

[5] Recent international relations scholarship has highlighted this dynamic in the context of general purpose technologies. Consider, for example, Ding and Dafoe, "Engines of Power."; Jeffrey Ding and Allan Dafoe, "The Logic of Strategic Assets: From Oil to AI," *Security Studies* 30:2 (2021): 182-212.

States can, of course, discover readily extractable resource deposits in a short timeframe and these can be used to sustain an expansion of the state's military and/or greater investment toward research and development, procurement, training, reforms, etc. to respectively attain quantitative and qualitative gains in their fielded military capabilities. In principle, economic change may therefore have a more direct mechanism through which to affect the fielded military capabilities of states. In practice, however, resources typically take time to extract particularly if new infrastructure is required and most resources must be converted into a useful product (whether through refinement or a manufacturing process to take that raw material into a component or final product) and cannot be used as armaments. Whether cotton, copper, diamonds, petroleum, rubber, or titanium, readily extractable resources can only serve as inputs into the economy and more generally contribute to an expansion of a state's military potential. And while a resource boom can plausibly create a major imbalance in the resources that states in a dyad can allocate toward their military capabilities, resource booms, whether the discovery of veins of silver or the discovery of oil deposits, tend to have lagged effects on fielded military capabilities of states given the need to extract the resources and thereafter convert military potential into fielded military capabilities.

For these reasons, economic change lacks direct mechanisms through which to affect the fielded military capabilities of states and does not lend itself to engendering capability shifts that are both sizeable and abrupt. This, to be clear, does not mean that economic changes are inconsequential for international politics and have no effects on fielded military capabilities. Rather, the scope of the claims being made in this dissertation only concerns the suitability to the role of a prime mover engendering

sizeable and abrupt shifts in fielded military capabilities. In this specific respect, economic change is an unsuitable candidate for the role of a prime mover.

2.2.3 Diplomatic Change

Alliances have long been a central focus of international relations scholarship and diplomatic change is therefore frequently cited as a source of shifts in the military capabilities of states.[6] In principle, alliances allow two or more states to collectively employ their military capabilities against one or more adversaries and alliances can, as such, affect military balances. It is, however, important to recognize that while alliances can make some states feel "more secure" while making other states feel "less secure," entering into an alliance does not increase a state's fielded military capabilities, at least not in the short run. Alliances only (re)aggregate existing military capabilities such that there is no net increase in military capabilities present in the international system following the establishment of an alliance.[7]

Assuming that States A and B enter into an alliance directed against State C, State C should not experience a decline in its military capabilities in the State A-State C and State B-State C dyads unless the State A-State B alliance heralded a fundamental reorientation of State C's relations with both State A and State B. If State C previously had adversarial relations with both State A and State B, then State C would have had

[6] Stephen M. Walt, *The Origins of Alliances* (Ithaca, NY: Cornell University Press, 1990)
[7] Granted, alliances can allow alliance members to field a greater combined military capability set than the sum of their individual (separate) capabilities, but this invariably requires time and effort and is not therefore an intrinsic quality of an alliance. The case of NATO is instructive in that it took decades for standardization and economies of scale to take effect in areas ranging from doctrine and technology development to training and logistics. Short of a commitment to longer-term integration of this sort, which is not a given, alliances only aggregate existing fielded military capabilities.

every reason to consider both of its adversaries' capabilities into account when assessing the adequacy of its fielded military capabilities (before the establishment of the State A-State B alliance). As a result, State C does not experience any decline in its fielded military capabilities following the establishment of an alliance between State A and State B. It is only when State C did not previously have adversarial relations with either State A or State B (or both states) that a State A-State B alliance affects State C's military capabilities and even in such a situation the new alliance does not introduce any net new military capabilities in State C's orbit, does not engender a sizeable and abrupt negative shift in State's C fielded military capabilities, and, no less importantly, the cause for State C's newfound insecurity is the rupture in its relations with State A/State B, and not the binding of State A and State B's military capabilities per se.

In principle, alliances allow states to expand their fielded military capabilities by collaborating and pooling resources with their allies. While states can enter into an alliance very quickly, it takes time for the new allies to integrate capabilities (by training with one another, etc.), standardize equipment, coordinate military planning, coordinate new research and development or procurement efforts, etc. Stated differently, it takes time – often years – for an alliance to be more than a simple aggregation of extant capabilities and most alliances have historically been of far shorter duration and characterized by far lower levels of integration than the North Atlantic Treaty Organization, which offers a template for what an alliance can achieve through many years of coordination while also serving as a cautionary tale of the challenges and limited benefits of high levels of coordination. Consequently, not only do alliances not independently engender any net increase of military capabilities in the international

system, but the effects of alliances on fielded military capabilities tend to be lagged. For these reasons, diplomatic change is not a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

2.2.4 Political Change

Although international relations scholarship has historically privileged inter-state political dynamics (i.e., diplomacy), recent scholarship increasingly examines the effects of domestic politics, broadly defined, on international politics.[8] As a result, various strands of scholarship highlight the ways through which political changes within states can contribute to shifts in the military power of states. Insofar as the perceived interests and policy preferences of states and, as such, state behavior, emanate from internal political factors, domestic political changes, whether elections, successions, or revolutions, can have a variety of significant effects on the threat perceptions of other states but a shift in fielded military capabilities is not necessarily one of these reasons for this. The new political leadership in State A, which, for example, recently experienced significant domestic political change, may pursue different policies, including an active pursuit of increased military capabilities, and this can challenge the assumptions that State B (and other states) made about its security, but there is no direct effect – or indeed any immediate effect – of such political changes on the State A-State B military balance other than a possibly increased salience of existing fielded military capabilities.

[8] Deborah D. Avant, *Political Institutions and Military Change: Lessons from Peripheral Wars* (Ithaca, NY: Cornell University Press, 1994).; Jason Lyall, Divided Armies: Inequality and Battlefield Performance in Modern War (Princeton, NJ: Princeton University Press, 2020); On the effects of international politics on domestic politics, consider Seva Gunitsky, *Aftershocks: Great Powers and Domestic Reforms in the Twentieth Century* (Princeton, NJ: Princeton University Press, 2017)

Although political change encompasses the textbook example of a "revolutionary" development in the form of both peaceful and violent changes in political leadership including elections, only the leadership and policies of a state can change immediately, not its fielded military capabilities. Even if a new political leadership is bellicose toward another state, the new leaders do not have any greater fielded military capabilities at their immediate disposal than their predecessors. The new leadership can, of course, allocate additional resources to the military and so forth but the effects on the balance of fielded military capabilities will be lagged and, in any event, leaders cannot allocate resources that they do not have at their disposal. As a result, the new political leadership must either mortgage the future of its country – incur debt to fund the expansion of its military or, alternatively, prioritize resources to its military at the expense of its economy and thereby fail to invest in expanding its military potential – or must benefit from an exogenous source of new resources (in which case the effects attributed to political change would be coproduced with other unspecified exogenous mechanisms). If a new political leadership undertakes the former approach, its adversaries can determine that the gains in its fielded military capabilities may be unsustainable in the long run and possibly unsustainable even in the short run. If a new political leadership benefits from an exogenous injection of new resources, then the new resources, which are likely to take time – years – to extract and convert into an increase in fielded military capabilities, then economic changes, not political changes per se, are doing the heavy lifting. A state can, of course, benefit from an exogenous injection of resources in the form of military aid from another state but in such scenarios, diplomatic

change, not political change per se, would be doing the heavy lifting in explaining the increase in a state's fielded military capabilities.

While new leaders can increase the size of their military and/or increase the resources allocated to their military, recent scholarship highlights the diverse menu of options through which states can improve their military capabilities. New political leadership can, for example, undertake changes in civil-military relations, patterns of equality and inequality, and so forth and these can offer states an increase in their fielded military capabilities.[9] Such reforms cannot, however, be implemented overnight and logically have the greatest effects on a state's fielded military capabilities in contexts in which civil-military relations, inequality, etc. were poor and heavily debilitating to a state's military capabilities to begin with. Stated differently, the greater the problem, the greater the scope for improvement such that gains in capabilities can become very impressive in the eyes of an observer more as a result of base effects than a state's decipherment of a novel means to punch above its weight once it corrects its shortcomings.

For these reasons, the effects of political change on a state's fielded military capabilities are logically shaped by the extent to which the state in question, for want of a better phrase, punched below its weight prior to the change in political leadership or policy direction. Political change nevertheless encompasses the phenomena of political

[9] Consider, for example, Lyall, *Divided Armies*.; Caitlin Talmadge, *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes* (Ithaca, NY: Cornell University Press, 2015).; Peter D. Feaver, "Civil-Military Relations," *Annual Review of Political Science* 2 (1999): 211-241.; Peter D. Feaver, "Civil–Military Relations and Policy: A Sampling of a New Wave of Scholarship," *Journal of Strategic Studies* 40, No. 1-2 (2017): 325-342.; James Quinlivan, "Coup-Proofing: Its Practice and Consequences in the Middle East," *International Security* 24, No. 2 (1999): 131–165.; Risa A. Brooks, *Shaping Strategy: The Civil-Military Politics of Strategic Assessment* (Princeton, NJ: Princeton University Press, 2008).; Caitlin Talmadge, "Different Threats, Different Militaries: Explaining Organizational Practices in Authoritarian Armies," *Security Studies* 25, No. 1 (2016): 111-141.

revolutions in which a new political leadership attempts to fundamentally reshape its society, political system, economy, and, not least, its military. In principle, revolutionary regimes can be agents of major expansions in fielded military capabilities, time and resources permitting.[10] Revolutionary regimes can mobilize larger segments of a population for military service, thereby offering a mechanism through which political change can engender sizeable and abrupt shifts in fielded military capabilities that is distinct from demographic change. It is, however, important to recognize that the direct mechanism for this potential prime mover does not expand the pool of manpower available to a state that can volunteer and/or be forced into military service. Stated differently, revolutions do not increase the maximum potential manpower available to a country – there is no automatic increase in the population of physically capable "fighting age" persons – and, no less, untrained, unequipped, unled, and unsupplied manpower increasingly amounts to little in terms of fielded military capabilities in recent centuries (as opposed to earlier timeframes). Furthermore, a revolution does not necessarily expand a country's economy and every additional person "allocated" to the military is a person not being economically productive and contributing to a state's military potential. In a revolutionary state as in any other, it takes time to raise, train, equip, and sustain new military forces. At the same time, revolutionary states tend to experience political dynamics that can divide rather than solely unite populations – the newly enfranchised political winners logically coexist (absent mass slaughter and/or population displacement) with newly disenfranchised political losers from the prior political order – and can thereby also reduce fielded military capabilities even if there is a net gain.

[10] Consider, for example, Jasen Castillo, *Endurance and War: The National Sources of Military Cohesion* (Stanford, CA: Stanford University Press, 2014).

Political revolutions do not therefore necessarily result in increased fielded military capabilities, and it is problematic to make the small subset of revolutions that do so serve as an archetype for the military effects of political revolutions.

Political revolutions and elections primarily affect politics at the top of the political pyramid. Political change may, however, extend far beyond the leadership of a country and encompass issues such as regime type, civil-military relations, and inequality. International relations scholarship increasingly highlights such issues as non-material factors that can affect the military power and/or the military capabilities of states and this subset of political change may constitute a potential candidate for the role of a prime mover. As scholars have argued, a shift in regime type can facilitate gains in terms of military effectiveness,[11] but such dynamics encompass non-technological military changes – revisions to military organization, doctrine, training, and so forth – and such revisions take time to plan and implement. The performance of the Egyptian military in 1973 relative to 1967, which constitutes perhaps the archetype case of a change in regime type resulting in increased military capabilities,[12] highlights the limited scope and magnitude of capabilities gains possible without the time and resources required to close a large gap with an adversary that itself continued to improve its military capabilities in the interim. Similar dynamics apply to some of the most prominent cases of shifts in

[11] See, for example, Dan Reiter and Allan C. Stam. "Democracy and Battlefield Military Effectiveness," *Journal of Conflict Resolution* 42, No. 3 (1998): 259-277.; Dan Reiter and Allan C. Stam, *Democracies at War* (Princeton, NJ: Princeton University Press, 2002). For counterarguments, see Michael C. Desch, "Democracy and Victory: Why Regime Type Hardly Matters," *International Security* 27, No. 2 (2002): 5–47.; Downes, "How Smart and Tough Are Democracies?"

[12] Risa Brooks, "An Autocracy at War: Explaining Egypt's (in)Effectiveness in the 1967 and 1973 Arab-Israeli Wars" Security Studies 15, No. 3 (2006): 396-430.

civil-military relations that are associated with gains in military capabilities.[13] Both changes in regime type and changes in civil-military relations highlight the degree to which the effects of political change on fielded military capabilities are constrained by preexisting conditions. Simply put, states with poorly performing militaries require considerable time to redress their limitations. Similarly, preexisting conditions heavily shape the effects of inequality and shifts in inequality within states on fielded military capabilities. As with poor civil-military relations, military inequality does not emerge overnight and tends to reflect underlying factors. For this reason, it is also unlikely to change overnight in ways that practically engender sizeable and abrupt gains in a state's fielded military capabilities.

For these reasons, political change lacks direct mechanisms through which to affect the fielded military capabilities of states and does not lend itself to engendering capability shifts that are both sizeable and abrupt. This, to be clear, does not mean that political changes are inconsequential for international politics and have no effects on fielded military capabilities; the scope of the claims being made in this dissertation only concerns the suitability to the role of a prime mover engendering sizeable and abrupt shifts in fielded military capabilities. In this specific respect, political change is an unsuitable candidate for the role of a prime mover.

2.2.5 Non-Technological Military Change

International relations scholars have long concerned themselves with a range of non-technological military developments, whether an increase in the size of a military,

[13] Consider, for example, the case studies covered in Talmadge, *The Dictator's Army*.

changes in doctrine, changes in training, and so forth. Although non-technological changes in a state's military capabilities may appear to be an ideal candidate for the role of a prime mover, the issue is that such changes in military capabilities tend not to engender abrupt shifts in fielded military capabilities. An increase in the size of a military, for example, requires a state to not only allocate additional manpower to its military but also the resources required to train, equip, and sustain the additional manpower. Unless a state is sitting on a large reserve of serviceable equipment, has surplus training capacity and trained commissioned and non-commissioned officers available to lead freshly raised forces, and has in place the requisite logistical capacity to sustain its expanded military, the effects of military expansion tend to be lagged and non-technological military change does not, as such, lend itself to engendering shifts in fielded capabilities that are both sizeable and abrupt.

There are, of course, a range of non-technological military changes beyond the mere expansion of a military in terms of numbers. Military reforms in the areas of organization, tactics, and doctrine, for example, constitute efforts toward improving a military's capabilities in qualitative terms. Although such efforts can engender considerable capability gains to a state, these invariably have lagged effects given that it takes time – many months if not several years – to develop and, once ready, to implement across a military. Moreover, the effects of revised training only affect the personnel who undergo the (re)training, training capacity is invariably limited, and trained personnel do not remain in a military indefinitely so at some point training capacity is utilized simply to maintain force structure at a given training level rather than expand force structure or even increase the percentage of existing force structure trained to a given level. It

therefore takes time for the effects of revised training to affect most let alone the entirety of a state's military, and the effects of the new training (and new doctrine, new organizational practices, etc.) on the military balance are therefore not likely to be abrupt even if these are ultimately of greater consequence. Admittedly, even partial reforms can provide capability gains to a state, but it is important to recognize that adversaries can respond to another state's military reforms of its own and that adversaries may also independently undertake reforms of their own before a state's reforms are fully implemented – a state experiencing the impetus to undertake military reforms is likely to have responsive and adaptive adversaries that constitute a moving target. States in an adversarial dyad may also pursue divergent or convergent reforms that can mutually blunt the (net) effects of each side's reforms. Although non-technological military changes have direct mechanisms through which to engender shifts in the fielded military capabilities of states, non-technological military changes are unsuited to engendering shifts in fielded military capabilities that are both sizeable and abrupt and are not therefore a suitable candidate for the role of a prime mover.

International relations scholarship requires a prime mover to engender sizeable and abrupt shifts in the fielded military capabilities of states and the aforementioned sources of change regularly invoked in scholarship are not suitable for the role. Although international relations scholarship identifies a large number of causes of military power and factors that explain variation, both temporal and cross-state, in military power, most of these candidates for the role of prime mover lack a direct mechanism through which to affect the fielded military capabilities of states (i.e., these tend to affect the military potential of states and not fielded military capabilities) and/or have decidedly lagged

effects which is to say that these do not lend themselves to engendering shifts in fielded military capabilities that are both sizeable and abrupt. To be clear, this is not to say that these causes of, or factors affecting, military power are unimportant, but that most of these unsuitable candidates for the role of a prime mover have considerable implications for international politics that tend to gradually transpire. As a result, other states have more time to respond through efforts to blunt if not negate the effects of a non-zero shift in a military balance, and the military effects of these changes are therefore more likely to be moderated, and as such more limited and potentially inadequate for the purposes of initiating a security dilemma like dynamic or similar.

Having ruled out demographic, economic, diplomatic, political, and (non-technological) military change as suitable candidates for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states, we are left with one remaining candidate, technological change, or advances in military technology. As previously explained, there are plausible direct mechanisms through which advances in military technology can affect the fielded military capabilities of states, at least in a purely monadic manner. The question remains whether advances in military technology can engender shifts in fielded military capabilities that are both sizeable and abrupt.

To address the question of whether technological change is a suitable candidate for the role of a prime mover, I will first develop a multi-faceted (i.e., both monadic and dyadic) analytical framework in the following chapter that reflects upon how military technologies are developed, tested, manufactured, deployed, and employed in times of both peace and war against adversaries that have incentives to undertake efforts aimed at

blunting if not negating the shifts in fielded military capabilities engendered by technological change. This analytical framework will allow us to pursue answers to two questions concerning the suitability of technological change for the role of a prime mover of military capabilities. First, is there a direct mechanism through which change in this area – technological change – can engender shifts in fielded military capabilities? Second, does change in this area – technological change – lend itself to driving shifts in fielded military capabilities that are both sizeable and abrupt in terms of effects? After developing an analytical framework in Chapter Three, I will then turn to two detailed longitudinal case studies, each composed of several chronologically separated chapters, in Part Two and Part Three of this dissertation to discern whether the analytical framework accords with the historical record or if technological change has tended to engender sizeable and abrupt shifts in the fielded military capabilities of states. Specifically, I will examine the implications of the advent and evolution of submarine technology over the ca. 1800 to 1945 timeframe and the implications of the advent of uncrewed aircraft technology over the ca. 1910 to 2020 timeframe. These case studies will serve as plausibility probes for the analytical framework across time and space and together help lay the groundwork for future research on the drivers and implications of shifts in fielded military capabilities for international politics.

Chapter 3: An Analytical Framework for Assessing the Military Effects of Technological Change

This chapter develops a simplified analytical framework through which to discern whether technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded capabilities of states. As with the prior exercise of ruling out other candidates for the role of a prime mover of fielded capabilities, this chapter sets out to answer two questions about the relationship between technological change and military capabilities. First, is there a direct mechanism through which change in this area – technological change – can engender shifts in fielded capabilities? Second, does change in this area – technological change – lend itself to driving shifts in fielded military capabilities that are both sizeable and abrupt in terms of effects? This chapter will attempt to answer these questions while also establishing baselines for examining the implications of real-world technological advances that will be examined in later chapters. The analytical framework developed here will explain that the effects of technological change on fielded military capabilities are moderated by the ways in which new military technologies are developed, tested, manufactured, and deployed in times of both peace and war in a world in which responsive materially matched adversaries can attempt to adopt, counter, and offset a new technology with the aim of blunting if not negating its effects on the military balance. Because this chapter is an analytical exercise and does not examine real-world cases, it will be light on specific empirics and does not include a case selection methodology. Instead, this chapter first explicates a series of analytical assumptions and tenets and then in turn examines the

monadic and dyadic mechanisms through which advances in military technology can engender shifts in the fielded military capabilities of states.

3.1. Analytical Assumptions

Before proceeding, it is important to explicate a series of assumptions upon which the analytical framework is predicated.

3.1.1. Assumption One: States have finite resources, and these must be prudently allocated.

In a hypothetical world in which states can draw upon infinite resources, states would have an unbounded menu of options through which to maintain their military capabilities relative to other states. In the real world, all states, big and small, have temporally finite resources at their disposal and must prudently allocate these resources in pursuit of various – and often competing – objectives. Allocation can, in general terms, go toward consumption or toward investment, and the fact that resources are finite means that every "unit" of resources allocated to sustain or improve fielded military capabilities is one "unit" of resources that cannot be allocated toward some other purposes. For this reason, all states encounter a complex array of allocative trade-offs that need to be considered when evaluating whether technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in fielded military capabilities. For example, while resource allocation toward infrastructure and education can increase a state's military potential over time, such efforts tend to have negligible effects on the

immediate and near-term fielded capabilities of a state largely for want of a direct mechanism but also because of the typically lagged effects of an increase in a state's military potential on its fielded capabilities. At the same time, a state undertaking maximal investment toward improving fielded military capabilities would necessarily result in limited investment (if any) toward expanding its military potential in the long run. If international politics were like a sports game in which the contestants have only to husband and allocate their resources within a fixed – and predetermined – timeframe, then the (inter-temporal) trade-offs associated with resource allocation and the risks and uncertainties involved would not be so great. And if resources were not finite, then states could allocate without consequence or opportunity cost. In the real world, however, states must prudently allocate their finite resources, and this should be the starting assumption when examining whether technological change is suited to the role of a prime mover engendering shifts in fielded capabilities that are both sizeable and abrupt.

While states must decide between (A) allocating resources toward military or non-military purposes and (B) allocating resources toward increasing their military potential or their fielded military capabilities, resource allocation is not an abstract exercise. Even when states allocate resources with the aim of improving their fielded military capabilities, states must prudently determine specifically where and toward what specifically they should allocate their finite resources. For example, states must decide between investing in maintaining (or expanding) force structure (i.e., manpower), funding maintenance activities, funding training, funding readiness, funding the procurement of new equipment, and so forth. A state that maximally allocates toward force structure at the cost of maintenance and readiness, for example, may be neither ready to engage in

combat operations on short notice nor be capable of sustaining combat operations for long thereafter even while appearing formidable on paper. Similarly, a state that maximally allocates toward force structure (i.e., manpower) and procuring equipment at the cost of training may have very limited fielded capabilities in terms of the missions that it can successfully undertake in a given time and place. Although all states encounter trade-offs and opportunity costs when allocating their finite resources toward their fielded military capabilities, the specifics are consequential and should not escape consideration in an analysis focusing on the development, diffusion, and employment of military technologies.

3.1.2 Assumption Two: There is cross-national variance in state resources in the real world, but it is analytically productive to assume that states are materially matched to control for this major confounding variable.

Although all states have finite resources at their disposal and consequently encounter allocative trade-offs, there can be considerable cross-national variation in the resources available to states and this is a major confounding variable when attempting to discern whether technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in fielded military capabilities. State A, for example, may successfully develop a potentially effectual military technology *x* but may nevertheless lack the resources to manufacture and deploy technology *x* in sizeable numbers. In such a situation, State A would likely only benefit from small gains in its fielded military capabilities as a result of its development and deployment of technology *x* and the limited magnitude of effects on the State A-State B military balance could

potentially allow State B, State A's adversary, to more readily blunt if not negate the effects of technological change on its fielded capabilities by undertaking efforts to adopt, counter, and offset technology *x*. In fact, with State A being unable to make the most of the technological advance it introduced to the State A-State B military balance, State B may be able to both adopt the new technology *x*, thereby accruing the capability gains that were initially monopolized by State A, as well as counter the new technology provided that it has the resources to simultaneously pursue both approaches. In such a situation, State A could find itself worse off than it began as a result of its pursuit of the new technology *x* (assuming that State B was not already pursuing technology *x* and would not have otherwise done so for its own reasons). Alternatively, State B could first introduce technology *x*, and State A could attempt to counter it but lack the resources to deploy (technological) countermeasures in adequate numbers to blunt let alone negate the effects of technology *x* on the State A-State B military balance. In these examples, the military effects of technological change on international politics would be strongly shaped by disparities in resource availability which is to say extra-technological factors.

Given that significant cross-national variance in the resources available to states can, in effect, distort the effects of technological change on the balance of fielded military capabilities in a (real-world) state dyad and thereby serve as a major confounding variable for analysis, the analytical framework developed in this chapter is purposefully written *as if* states are in fact materially matched and balanced in terms of resource availability. This facilitates a more direct examination of the effects of technological change on the fielded military capabilities of states by holding other variables constant. After all, it is important to discern *whether* states can blunt or negate the shifts in fielded

military capabilities engendered by technological change independently of whether they have the resources to do so. If there is no scope for a state to blunt or possibly negate a (negative) shift in its fielded military capabilities engendered by technological change, then any real-world resource disparities are immaterial, and the state in question will incur the (unmoderated) effects of technological change in its entirety. If, however, there is scope for a state to blunt or possibly negate the effects of technological change on its fielded capabilities, then the effects of technological change on fielded military capabilities can be moderated such that the military implications for international politics may be more limited.

Although the reality of cross-national variance in resource availability cannot be entirely ignored, the analytical framework developed in this chapter suggests that for as long as there is broad, not necessarily exact, parity in the resources available to members of a state dyad, it becomes very difficult, if not close to impossible, for one state in the dyad to both obtain *and* sustain a significant advantage in fielded military capabilities by developing and deploying new military technologies. As such, if one state both obtains and sustains a significant advantage in its fielded military capabilities by developing and deploying new military technologies, this may reflect upon extra-technological factors such as resource disparities, among other factors. All things equal, states in materially balanced dyads are well positioned to blunt if not negate the effects of technological change on their fielded capabilities by adopting, countering, and/or offsetting a new technology. If nothing else, this analytical assumption of resource equality allows us to establish baselines through which to compare real-world dynamics and thereby discern

the “marginal effects,” as it were, of resource disparities on the implications of technological change in future research.

3.1.3 Assumption Three: States are motivated actors that can respond to prospective and realized shifts in fielded military capabilities with the aim of blunting if not possibly negating the effects thereof.

While the availability of resources enables states to adopt, counter, or offset a new military technology, states are led by human actors with agency who must first decide whether they want to respond to a realized or projected (negative) shift in their fielded military capabilities. States may choose not to respond to a material shift in their capabilities for any number of reasons beyond the potential inadequacy of their finite resources and associated allocative trade-offs. For example, State A may not be concerned about the new technology *x* if it is introduced (i.e., first developed and deployed) by State C, a state with which it has friendly relations. In such a situation, State A’s threat perceptions could be dampened by non-military and non-technological considerations such that State A is essentially unresponsive to the introduction of technology *x* even as State A experiences a relative decline in its fielded military capabilities vis-à-vis State C, and this may give the appearance that State A is an unmotivated actor. State A may, however, be (more) responsive if technology *x* is introduced by State B, which State A views as an adversary, or if State B adopts, counters, or offsets technology *x* in response to State C’s introduction of the new technology and its effects on the State B-State C military balance. In such situations, State A would appear to be a more motivated actor but none of this is to say that there is

necessarily any variance in the feasibility of State A adopting, countering, or offsetting technology *x* between these notional scenarios.

In the real world, states tend not to encounter a dichotomous choice between responding and not responding to the introduction of a new military technology by one or more states. States can, and often do, hedge their bets, and assume some level of risk in doing so but nevertheless better position themselves to adopt, counter, or offset a technological advance when threat perceptions are held to justify the resource allocation. These hedges may or may not pay off and state decisions can be overtaken by political and diplomatic events but the implication of this brief analytical exercise is that the choice of a state not to undertake efforts, or not undertake maximal efforts, to blunt or possibly negate the effects of technological change on its fielded military capabilities does not necessarily mean that it *could not* have done so through adopting, countering, or offsetting. To address this potentially confounding matter for the present purposes, the analytical framework developed in this chapter assumes that states are highly motivated actors attuned to prospective and realized shifts in their military standing relative to other states. Stated differently, the analytical framework development in this chapter assumes that states *actively* attempt to blunt if not negate the effects of technological change on their fielded military capabilities by adopting, countering, or offsetting the new technology. This helps to account for the possibility that states may decide to not attempt to blunt or negate the effects of technological change for reasons unrelated to the feasibility of doing so and for reasons unrelated to resource scarcity or allocative trade-offs.

3.1.4 Assumption Four: States seek out information on other state's military capabilities and technological efforts and use this information with the aim of blunting if not negating the effects of technological change (among other things).

Presuming adequate resource availability, an attentive and motivated state can respond to the introduction of new military technology by other states through the undertaking of efforts to adopt, counter, and/or offset said new technology so as to blunt if not negate the effects of technological change on its fielded military capabilities. To do so, states require *some information* about other states' military capabilities and technological efforts. Without any knowledge of what other states, particularly adversaries, are up to, a state is poorly positioned to prudently allocate its finite resources to maintain its military standing in relation to other states. Beyond possibly seeking to introduce – develop and deploy – new military technologies on its own accord, it would lack a "target" technology or capability set against which to undertake efforts toward adoption, countering, or offsetting. A state can, of course, assume that an adversary will have military capabilities similar and perhaps equal to its own (i.e., a form of mirror imaging). Although this may set a higher than necessary bar in some real-world cases, a state may nevertheless find itself surprised when encountering other adversaries with heterogeneous capability sets for which such assumptions do not hold.

To mitigate against such surprises, the leaders of states (polities) have historically sought out information from formal and informal spies and informants to develop a better understanding of other states including how the military capabilities and technological pursuits of these states compare with that of their own. Through such endeavors, states can obtain information about the present condition of another state's military capabilities

– its strengths and its weaknesses – and the technologies that it is pursuing and fielding. Importantly, this information need not be perfect or complete. An understanding of the strengths and weaknesses of State B's military capabilities can, for example, present State A with options as to which types of technologies it (State A) may benefit from introducing as well as which types of technologies State B may itself attempt to introduce assuming, of course, that State B shares State A's perception of its own strengths and weaknesses in relation to those of State A. And if State A learns of a new technology that State B is developing and intends to eventually deploy, then State A is better equipped to make informed decisions about how to allocate its finite resources in order to blunt if not possibly negate the effects of technological change on the State A-State B military balance.

In the real world, there can be tremendous variation – cross-national, temporal, and otherwise – in the extent to which one state successfully assesses another state's military capabilities and which technologies another state is pursuing. At times, states have failed to even detect the development of a new technology by an adversary and have therefore been slow – and sometimes far too late – to initiate efforts to blunt or negate the effects of the technological change on their fielded military capabilities.[14] In such cases, the effects of technological change may be unmoderated, or at least less moderated, by otherwise possible state actions and can as such have greater implications for international politics. This, however, does not necessarily mean that the effects of technological change could not have been blunted or possibly negated through efforts toward adopting, countering, and/or offsetting a new technology. To account for the

[14] Thomas Mahnken, *Uncovering Ways of War: U.S. Intelligence and Foreign Military Innovation, 1918-1941* (Ithaca, NY: Cornell University Press, 2002).

possibility that motivated states may fail to respond to the technological pursuits undertaken by a materially matched adversary, this analytical framework assumes that states actively seek out information on other states' military capabilities and technological pursuits. The information obtained through such efforts can be incomplete and imperfect but can nevertheless better position states to make more informed decisions as to how they should allocate their finite resources and how they may be able to blunt if not negate the effects of technological change on their fielded military capabilities.

3.2 Analytical Tenets

Accompanying the aforementioned analytical assumptions are what are best characterized as several analytical tenets. Explicating these tenets sheds light on how this analytical framework has been developed and why it comes to its (preliminary) findings. This is important as this chapter is a purely analytical exercise that will be followed up by a series of case studies that will test whether this analytical framework accords with the historical record as to whether advances in military technology have (historically) engendered sizeable and abrupt shifts in the fielded military capabilities of states.

3.2.1 Tenet One: The Speed of Technological Change is Irrelevant to the Analytical Task at Hand.

First, whereas many writings dealing with advances in military technology emphasize the *speed* of technological change (often in abstract and absolute terms), this analytical framework posits that the speed of technological change is essentially irrelevant to the

analytical task at hand.[15] In itself, technological change lacks a direct mechanism through which to affect the primary outcome of interest for international relations scholars (and many policymakers), which concerns shifts – changes – in the fielded military capabilities of states in relation to one another. Simply put, the effects of technological change on the fielded military capabilities of states are *zero* unless and until a new technology is put into production and deployed in militarily significant numbers. The speed of technological change does not, therefore, have any significance independent of the process of translating, or converting, technological advances into fielded military capabilities. If anything, rapid technological change creates challenges for states that serve to further moderate the military effects of technological change as necessarily resource-constrained states face difficult decisions as to which iteration of a rapidly changing technology to commit to production.

As explained in the previous chapter, it is analytically important to examine both the pace at which shifts in fielded military capabilities unfold and the magnitude of such shifts given that the metric of ultimate importance is the magnitude of shifts in fielded military capabilities per unit of time (i.e., the "speed" of capability shifts). All things equal, the more time that is available to a state before a technological change engenders a significant shift in the military balance, the more time a state has to adopt, counter, and offset the technological advance and thereby attempt to blunt if not negate the effects on its military capabilities. The practical implication of this consideration is that a sound analytical framework needs to view technological change as a process beginning but not

[15] Commonly expressed in terms of one or more metrics often associated with performance and/or cost.

ending with developmental efforts rather than viewing technological change as an outcome with intrinsic effects that unfold once development is completed.

3.2.2 Tenet Two: A Sound Analysis of the Implications of Advances in Military Technology Must be Longitudinal in its Approach.

Second, given the importance of viewing the development of new military technologies as processes rather than outcomes to be examined after the fact, a sound analytical framework must be longitudinal in its approach. Unless an analysis of the implications of technological change for international politics is longitudinal – begins with the origins of an area of technology and thereafter tracks its evolution from prototype to deployed equipment and thereafter examines its effects and co-evolution alongside countermeasures and other technologies – observers are susceptible to examining the effects of new technology only when it is, on a problematical post hoc basis, thought to have mattered or have been effectual in the form of highly curated and unrepresentative snapshots masquerading as case studies. Doing so effectively amounts to selecting cases on the outcome of interest and constitutes an epistemologically fraught approach through which to extrapolate generalizable inferences about the military effects of technological change. Although necessarily more complex, a longitudinal analytical framework avoids such problems by tracing the pace and magnitude at which shifts in fielded military capabilities transpire over time and thereby capturing periods in which the effects of technological change were greater and lesser. It begins with the origins of a new technology and does not stop after the first indications of a technology's effects but rather continues to track the implications of successive technological advances in a given area

without setting an arbitrary endpoint for analysis. In so doing, a longitudinal analytical framework can attempt to reconstruct the implications of technological change as these were encountered by (then-contemporaneous) human actors, including policymakers, in chronological real-time who did not benefit from foresight.

3.2.3 Tenet Three: A productive analysis must holistically examine a representative set of technologies.

Third, while a longitudinal framework helps to guard against the possibility that an analysis places a disproportionate (if not exclusive) emphasis on periods of temporally atypical variance in the effects of technological change, a productive analysis of the implications of advances in military technology for international politics must also holistically examine a representative set of technologies to account for potential variance in the pace and magnitude of shifts in capabilities across areas of military technology. Failing to do so leaves an analysis unequipped to answer the question of whether its findings are generalizable or simply reflective of a narrow and broadly unrepresentative area of technological development.

A longitudinal framework for analysis can, for example, indicate that some technologies experience periodic "growth spurts" but afterward return to modest rates of performance growth. What a longitudinal analysis cannot independently rule out is whether the observed growth spurt is itself typical for most military technologies. Technology development in the areas of software and electronics, for example, has historically been characterized by high rates of performance growth which, notwithstanding considerable lag times before new technologies feature in the equipment

available to frontline military forces in militarily significant numbers, have nevertheless translated into an ongoing series of increasingly capable defense electronics such as aircraft radar that can, in principle, engender major gains (shifts) to an adopting state's fielded military capabilities. Without a holistic analysis of a representative set of technologies, observers may, however, overlook how aircraft maneuverability, which is primarily a function of airframe design and propulsion technology, has mostly plateaued in recent decades with modest improvements obtained at a considerable expense even in the context of significant growth in aircraft radar technology. Although an observer may, for example, nevertheless contend that the advances in aircraft radar performance outweigh the plateauing of performance growth in the area of aircraft maneuverability on fielded military capabilities, only a holistic analysis of a representative set of technologies can determine whether growth in aircraft radar performance, to continue with the example, is typical and extrapolatable across other areas of military technology.

3.2.4 Tenet Four: An analytical framework must consider both monadic and dyadic mechanisms.

Fourth, a sound analytical framework examining the effects of technological change on the fielded military capabilities of states must consider both monadic and dyadic mechanisms. By monadic mechanisms, I refer to the shifts in fielded military capabilities engendered by technological change independent of the actions taken and forgone by an adversary. Stated differently, a monadic analysis is conducted as if no adversary exists; it is only concerned with the development of a new technology, its manufacture, and the process of deploying the new technology. Focusing on monadic dynamics is analytically

productive because it allows for the establishment of baselines about the pace and absolute, or gross, magnitude of shifts in fielded military capabilities engendered by a technological advance. If the effects of advances in a given area of military technology are limited even without an adversary undertaking efforts to blunt if not negate the engendered capability gains, then the military effects of a given technological change are moderated, and the new technology in question may be of limited significance for international politics.

In the real world, many states for whom advances in military technology and shifts in military power are politically salient tend to view one or more states as potentially threatening and the state dyads in which shifts in fielded military capabilities are most salient are those which by definition involve at least one state that views another as its adversary. Whereas an analysis of monadic mechanisms disregards actions that can be taken and forgone by an adversary and thereby highlights the absolute gains in fielded military capabilities that may be engendered by technological change, an analysis of dyadic mechanisms focuses on the efforts that other states can take to blunt if not negate the effects of a technological advance by adopting, countering, and/or offsetting a new technology (these are not mutually exclusive). By doing so, a dyadic analysis offers insight into the net effects on the balance of fielded military capabilities between two states and it is the net (relative) effects resulting from competitive interaction, not the absolute (gross) effects of a single-actor race, which are of ultimate consequence for international politics.

Although the distinction between monadic and dyadic mechanisms is inherently artificial – no real-world state experiences one independent of the other – it is

nevertheless analytically productive, particularly for the purposes of developing a framework that attempts to account for confounding variables of which there are many. A monadic conception of the military effects of technological change emphasizes the pace of change in fielded military capabilities as well as the absolute magnitude of effects while a dyadic conception of effects emphasizes the net magnitude of shifts engendered in fielded military capabilities in a dyad featuring responsive state actors. In the real world, and in the examination of real-world cases in the chapters that follow this present chapter, monadic and dyadic effects are of course intertwined just as the pace and magnitude of shifts in fielded military capabilities cannot be independently understood without reference to the other.

3.3 A Monadic Framework for Assessing the Pace and Magnitude of Shifts in Fielded Military Capabilities

As explained earlier, advances in military technology have *zero* effects on fielded military capabilities unless and until such technologies enter frontline service. For this to happen, a new technology must be developed, it must be tested, and progressively refined based on the results of testing, it must be made ready for production, it must be produced in sizeable numbers, preparations must be made for its entry into service (training, logistics, doctrine, etc.), and only then does a new technology enter frontline service and *begin* to affect the fielded military capabilities of a state. Although entry into frontline service reflects the moment at which a new technology *begins* to have a material effect on a military balance, it is important to recognize that the magnitude of the effects (over time) is a function of the prevalence of the technology (i.e., how many units or specimens

are in service), the fraction of a state's capability set that these constitute, the ways that the new technology is employed in an invariably hybrid force structure alongside (older) existing technologies (at least in the interim), and, not least, how one state organizes and plans to employ the new technology (i.e., force employment, doctrine) vis-à-vis its adversaries.

Even without an adversary undertaking any efforts to blunt or negate the effects of the new technology at any of the above stages, these monadic mechanisms moderate the effects of technological change on the fielded capabilities of states and suggest that technological change is not a suitable candidate for the role of a prime mover engendering capability shifts that are both sizeable and abrupt. To be clear, this does not mean that technological change has zero effects on military balances. The argument is simply that converting ideas for a new military technology into a production-ready design that is then deployed in sizeable numbers tends to take considerable time – often many years[16] – and this inherently moderates the military effects of technological change for want of abruptness. To understand how this is so, it is productive to divide the process of converting an idea for a new military technology to fielded military capabilities into four phases: development, testing and trials, manufacturing, and deployment. These monadic phases – in which no adversary state undertakes any efforts to affect the capability gains accrued to the state adopting a new technology – shape the pace of changes engendered in the fielded military capabilities of states and, as such, the magnitude of changes engendered per unit of time.

[16] This is an empirical claim that will be laid out in Part Two and Part Three of this dissertation project.

3.3.1 Phase One: Development

As far as international relations scholarship focuses on military technology, it tends to focus on the effects and implications of technological change and not the processes of how new military technologies come into being. As a result, scholarship largely examines military technology in terms of high-level political decision-making, organizational and bureaucratic politics, and intra-military debates at the intersection of technology and doctrine.[17] Stated differently, scholarship tends to examine the role of political leaders, senior bureaucrats, and senior military officers, most of whom are neither the source of the ideas that take material form as a specimen of military technology nor the individuals who facilitate the transformation from concept to deployable capability. New technologies are not developed in a vacuum, and although political leaders, bureaucrats, and military officers undoubtedly play a significant role worthy of scholarly examination, the effects engendered by a new technology are more directly shaped by the efforts of the individuals, and groupings of individuals, who conceptualize, develop, test, and manufacture new military technologies as well as those who operate and maintain it. After all, new military technologies have zero effects until these are deployed and new military technologies cannot be deployed until these are manufactured, and new

[17] See, for example, Owen Cote, "The Politics of Innovative Military Doctrine: The U.S. Navy and Fleet Ballistic Missiles" (Ph.D. dissertation, Political Science, Massachusetts Institute of Technology, 1996); Harvey M. Sapolsky, *The Polaris System Development: Bureaucratic and Programmatic Success in Government* (Cambridge, MA: Harvard University Press, 1972); Michael H. Armacost, *The Politics of Weapons Innovation: The Thor-Jupiter Controversy* (New York: Columbia University Press, 1969); Peter Dombrowski and Eugene Gholz, *Buying Military Transformation: Technological Innovation and the Defense Industry* (New York: Columbia University Press, 2006); Eugene Gholz, "Globalization, Systems Integration, and the Future of Great Power War," *Security Studies* 16, No. 4 (2007): 615-636.; Marc R. Devore, "Arms Production in the Global Village: Options for Adapting to Defense-Industrial Globalization," *Security Studies* 22, No. 3 (2013): 532-72.; Harvey M. Sapolsky, Benjamin Friedman, and Brendan R. Green, editors. *US Military Innovation Since the Cold War: Creation without Destruction* (New York: Routledge, 2009); Donald MacKenzie, *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance* (Cambridge, MA: MIT Press, 1990)

technologies cannot be manufactured until these are first developed. It is therefore only logical to begin with an examination of the mechanisms through which technological change can affect the fielded military capabilities of states with the development process.

3.3.2 Phase Two: Testing and Trials

Although iterative testing of components and component assemblies is invariably part and parcel of the developmental process, it is analytically productive to conceptualize it as an intermediate step between development and manufacturing – as testing and trials by the intended end user, the military. A design team undertakes testing with the aim of improving upon its technology toward a finalized design that is ready for production. As such, a design team will test individual components, component assemblies, and prototypes to address issues not limited to reliability, material fatigue, electrical faults, and bugs in software. Left to their own devices, design teams may not be particularly interested in putting their design through arduous tests and trials intended to identify its every weakness and how it fares in terms of both strengths and weaknesses relative to competing designs and substitutes. A design team is, after all, hardly an unbiased grouping of individuals regardless of whether the designers in question are in the employ of a state, employees of a profit-seeking enterprise, or more independent designers-turned-salespeople. With the market for military technology, particularly higher-end military technologies, being smaller than that of most civilian technologies and with states being the primary if not only potential customers of most military technologies (small arms have historically constituted a notable exception to this characterization),

design teams also tend to have a vested interest in having their designs selected by states particularly if there are finite resources to access and profits to be made.

For this reason, states and their militaries have historically put prototypes of new technologies through testing and trials of their own before selecting one or more designs and proceeding to allocate resources toward production. For this to take place, prototypes must first exist and be made available to a state so that the intended end user, the military, can carefully examine the new technology, identify its strengths and limitations, make a determination about the technology's prospects, and come to a conclusion about the overall suitability of the new technology (in its extant form) for its needs given the trade-offs in resource allocation and the potential availability of competing designs and substitutes. The military personnel[18] overseeing testing and trials may come to the view that a new technology offers relatively little compared to existing technology to justify the resource allocation – it may, for example, be more cost-effective to upgrade existing equipment. Alternatively, testing and trials may reveal a promising technology that requires further development before it is ready for frontline service, and this may be held to warrant the allocation of a military's resources toward continued development. Based on the findings of testing and trials, a military may put its procurement plans on hold until further advances are made in this area of technology, or it may decide that its finite resources are best allocated elsewhere and recommend terminating the military's funding of a design team. Alternatively, the military may determine that the new technology is precisely what they are looking for and may eagerly adopt the technology and push for the allocation of resources to initiate production.

[18] This is shorthand for uniformed members of a military, civilians employed by a military or the administrative body overseeing the military, and civilians brought in as outside experts/examiners, all of whom ordinarily reflect the interests of the military rather than the design team or manufacturer.

Without thorough testing and trials of the prototypes of a new technology, states would have to rely upon the veracity of claims made by design teams with a vested interest in securing additional funding for further development and/or the start of production. To be clear, a prototype of a new technology may work well enough in more controlled laboratory-like conditions and the issue is not necessarily that design teams lack scruples. Rather, militaries require rugged equipment that can reliably operate as intended on a sustained basis given that it is a matter of life or death for the military personnel operating the new technology in combat conditions. By having military personnel – rather than scientists, engineers, mechanics, and technicians from the design team – put a prototype through its paces, militaries are better positioned to discern whether the new technology will work as advertised in the roles and environments for which it is intended. Limitations and faults identified during military-run tests and trials may not spell the end of state support for the development of a new technology as these may be rectifiable once identified, and militaries may ultimately decide to accept the limitations of a new technology as a trade-off that accompanies its more desirable qualities. To facilitate thorough tests and trials, militaries typically hold design competitions for which one or more design submissions will typically receive state funding for continued development and/or a production contract. By doing so, militaries are exposed to a wider range of new technologies that are under development and are thereby better positioned to make informed decisions as to which technology to fund and/or procure.

Although military-run tests and trials, and thorough testing and trials more generally, are desirable for the aforementioned reasons, there is an important trade-off to

consider with implications for the question of whether technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt capability shifts in fielded capabilities. Simply put, thorough testing and trials take time to conduct and by definition force states to delay making procurement decisions until the tests and trials are completed and the results are reviewed. Military-run tests and trials may therefore improve the likelihood that a military prudently allocates its finite resources toward a selection of new technologies that are better positioned for success (i.e., positioned to deliver advertised capability gains), but this tends to come at the cost of moderating the effects of technological change on a state's fielded military capabilities by slowing down the transition from development to deployment. Of course, a state intent on realizing capability gains at the earliest possible opportunity, such as a state facing pressing wartime needs, can elect to forgo thorough testing and trials or a design competition and instead assume the risk that the new technology will not be as suitable for the intended roles as anticipated if and when it is tested in the crucible of combat. The new technology may not prove to be reliable in combat conditions, it may not be practical across the intended use cases, and it may prove to be readily susceptible to countermeasures – and it is possible that all these issues could have been identified through rigorous testing and potentially addressed through further research and development efforts. While there is no guarantee that military-run testing and trials will reveal all such limitations or that identified limitations can be rectified, states nevertheless encounter strong incentives to cautiously adopt new military technologies rather than recklessly assume considerable risk by pursuing every seemingly promising technology without first putting prototypes through thorough tests and trials.

3.3.3 Phase Three: Manufacturing

For a new technology existing in prototype form to have an appreciable effect on the fielded military capabilities of a state, it must be put into production so as to be available in larger numbers. Unless and until specimens are built and made available for militaries to deploy in sizeable numbers, even an otherwise promising prototype may have a negligible impact on a state's fielded military capabilities. Although states can (and have at times) put prototypes into frontline service, this has typically been a result of dire need or purposeful efforts to further test the equipment on a limited scale in real-world combat conditions. Prototypes are by their very nature essentially artisanal items that tend to be time-consuming, difficult, and expensive to build and operate (maintain). And because prototypes are invariably built in small numbers, the process for their manufacture does not benefit from economies of scale and the manufacturing methods that can reduce the marginal cost of production tend to have high fixed start-up costs that the production of a small number of prototypes cannot justify. As such, the manufacturing process is distinct from that of the development and testing of a new military technology even though there is some overlap.

Many writings on the implications of advances in military technology for international politics overlook the manufacturing process and essentially take it for granted.[19] This is problematic in two key respects. First, unless a single specimen of a new technology engenders a sizeable shift in the fielded military capabilities of a state, which is suspect, the number of specimens of the new technology in existence (and in the hands of trained operators) affects the magnitude of capability gains even in a purely

[19] Something best demonstrated by the paucity of international relations scholarship dealing with the manufacturing process.

monadic framework (which assumes that other states do nothing to blunt or negate the effects of such capability gains). Second, the manufacturing process serves as a bridge through which the development of a new technology can translate into a shift in fielded military capabilities. As a result, the manufacturing process is a bottleneck that regulates the throughput of the magnitude of shifts in fielded capabilities that a technological advance can engender per unit of time. For these reasons, both the pace and the magnitude of shifts in fielded military capabilities are heavily moderated by the manufacturing phase.

In principle, the manufacturing phase is straightforward. Having completed (i.e., frozen) the design of a new technology, factories are given blueprints and instructions, directed to begin production, and start production after sourcing suitable facilities, raw materials, tooling, and personnel. In practice, however, manufacturing is a complex process requiring the allocation of finite resources and comes with risks including the possibility that production will be delayed or that there are problems in the production process that undermine the performance of the new technology once it is deployed by the military. To appreciate the nature of this bottleneck on the shifts in fielded military capabilities engendered by the adoption of a new technology, it is productive to briefly consider some aspects of the manufacturing process.

In name, a state may want to produce a single new type of equipment, technology *x*. In reality, technology *x* is likely composed of a large number – perhaps many thousands – of distinct components. Although a single facility may handle the final assembly of the new technology and as such "build" specimens of the new technology, the production of components will invariably take place in multiple facilities each of

which constitutes a distinct cog in the production process and many of which will have distinct production equipment and can therefore manufacture some components while being wholly unsuited to producing others. Once built, these distinct components are typically sent to one or more facilities that handle final assembly, and the completed specimen of the technology is thereafter transferred to the end user, the military. There is, as such, not one bottleneck in the manufacturing process but many, and the final assembly facility's throughput can only be as high as the least reliable cog in the production process allows. Delays in shipments, failures in quality control, shortages in material inputs, and so forth at any of these component-producing facilities constituting the supply chain can reduce the output of the final assembly facility and thereby decrease the rate at which specimens of the new technology can be completed and transferred to the military.

Establishing the production process and putting in place the necessary supply chain is a complex endeavor that contributes to the lagged effects of technological change on fielded military capabilities. Although planning for the start of production tends to begin when a new technology is at an advanced state of development, the individuals responsible for the production process must wait for the design to be frozen (i.e., for the production model to be fully specified) before they put in orders for raw materials and components and allocate finite production capacity, tooling, and trained personnel. If the new technology requires novel production methods and tooling, then the manufacturing process may encounter delays until these are brought online. The manufacturing process can, of course, begin with less delay if no new production facilities are required and if existing tooling and production methods are suitable for the tasks at hand. Even so,

production output will start from zero and the individuals working on the various component production lines will need to accumulate experience so that production can be scaled upward toward maximum output levels. This process of scaling up production takes time and can also reveal a variety of issues with the new technology. The start of production can, for example, reveal defects that need to be addressed, such as a weakness in the grade of a material selected for a key component that may have been adequate during the rather artisanal manufacture of prototypes but not large-scale production. The individuals involved in the production process may also identify ways in which a design can be simplified to reduce the cost of production, the amount of materials required, and the number of hours of work that go into producing a component.

While scaling up production refers to increasing the production rate toward the maximum level attainable with a given final assembly facility, its supply chain, and resource allocations thereto, there are constraints on the maximum production rate that a state can hope to realize. Simply put, no state has the resources to undertake the production of all military technologies, both new and existing, at the maximum possible level. Instead, states must allocate their temporally finite production capacity, resource inputs, and, in market economies, fiscal outlays across a variety of military (and civilian) purposes. Production capacity, which encapsulates suitable equipment/facilities, suitable personnel, suitable resource/material inputs, and so forth, is expensive to establish and sustain and imposes fixed costs that, by definition, do not vary with the rate of production (typically up to an inflection point beyond which further fixed costs arise).

To put a new technology into production, a state can use existing production capacity (with or without modifications thereto), it can establish new production capacity,

or it can utilize some combination thereof. All things equal, the more novel a new military technology and the greater the performance gains on offer relative to existing military technology, the more likely it is that new production facilities and/or new tooling will be required. To rationalize the resource allocation toward the fixed costs required to initiate the production process, states face two competing objectives. First, states have incentives to achieve economies of scale to bring down the marginal cost of production. States do not, however, require an infinite number of a given new military technology, and economies of scale tend not to scale indefinitely; beyond a certain production rate, a state will have to cover additional fixed costs to expand production capacity, and this will result in an inflection point (increase) in the marginal cost of production, all things equal. To amortize the initial fixed cost of starting the production process (and any subsequent expansions of production capacity), states encounter a second structural incentive to extend a production run for as long as feasible. Here states again encounter the reality that their resources are finite and that they do not need – and in any event cannot afford to procure, sustain, or operate – an infinite quantity of a new technology. Compounding the challenges faced by states, these competing objectives exist in a world in which industrial capacity, whether physical structures, tooling therein, or trained personnel, is both finite and not perfectly fungible such that expensive industrial capacity developed for the production of technology *x* may be wholly unsuited to the production of technology *y* and so forth.

The practical effect of these competing pressures is that absent conditions of dire need, most new military technologies tend to have protracted – that is, deliberately drawn out – production runs, and this establishes the overarching bottleneck in the production

process irrespective of the maximum industrial capacity of a state. In a world of infinite resources, a state could establish such production capacity as to produce the entire production run of a new technology "in one go" (i.e., in one batch). With finite resources at hand, however, a state must prudently allocate so as to acquire *z* quantity of the manufactured specimens of new technology per unit of time and must do so by assessing both the marginal cost of production and the fixed cost associated with establishing or expanding production capacity. For such reasons, the manufacturing process plays an important role in moderating – curtailing – the pace and magnitude of shifts in fielded military capabilities engendered by technological change irrespective of the abstracted military implications of technological change.

3.3.4 Phase Four: Deployment

In principle, the effects of technological change on fielded military capabilities begin to accrue once completed specimens of a new technology leave a final assembly facility and are transferred to military personnel in a field unit. In practice, deployment is itself a distinct and often overlooked process, one that also tends to be protracted and associated with a variety of challenges and risks that tend to further moderate the effects of technological change on the fielded military capabilities of states even without the existence of an adversary (i.e., in a purely monadic analysis).

Although the efforts that need to be undertaken to deploy a new technology can vary considerably, militaries tend to need to create new training manuals and update their field manuals and doctrine so that personnel are trained to a consistent standard as to how a new technology should be operated and maintained on the basis of its strengths and its

limitations. The officers in command of military formations from platoon lieutenants to generals in the national headquarters need to learn about the new technology if the forces under their command are to effectively employ it and this may require the allocation of additional resources toward training and perhaps the procurement of other forms of equipment. For example, a new tank design may offer substantial improvements in terms of armor protection and firepower at the cost of increased weight, and this may result in increased logistical requirements such as new trucks with trailers capable of hauling the new tank at a desired speed, new armored recovery vehicles with adequate towing capacities, and new bridging equipment, all of which in turn require new training manuals and updated doctrine to be in place (alongside commensurate resource allocations toward the procurement and the sustainment thereof). Without making such material and non-material accommodations for a new technology, the real-world capability gains engendered by a technological advance will likely be less than anticipated and may even be negative (i.e., result in a net capability loss) if the infrastructure, human and physical, and support capabilities are not in place for a new technology entering frontline service.

Given that the production process constitutes a bottleneck and that resource-limited states are likely to have hybrid force structures comprising some combination of new and existing technology for some time until specimens of a new technology fully replace earlier designs, military personnel including their commanders must appreciate the different requirements for new and existing technologies and make suitable adjustments to operational planning and doctrine. The degree to which militaries

successfully do so when transitioning to a new military technology determines the real-world magnitude of shifts in fielded capabilities engendered by technological change.

If militaries simply employ a new technology as if it were the old/existing technology, then the nominal capability gains on offer may be unrealized. For example, a stabilized gun on a new tank may in principle allow tank crews to engage targets more accurately when the vehicle is moving, and this can have a decisive effect in a combat engagement. The effects – the capability gains – are not, however, automatically endowed through the deployment of the new technology as operators must be adequately trained to take advantage of the stabilization system, and maintenance personnel must be trained on how to keep the system serviceable. Otherwise, the new technology may have negligible effects and may in fact have negative effects on the fielded military capabilities of a state if the stabilization system falls into a state of disrepair or if it is incorrectly used. And if no tangible capability gains are obtained, then a state would have wasted scarce resources that could have been allocated elsewhere. Officers commanding military formations and those responsible for operational planning and doctrinal decisions encounter strong incentives to be considerate of such matters and ensure the allocation of suitable resources even if it comes at the cost of reducing the production run of the new technology. Otherwise, these officers, some of whom inform state leaders of the fielded capabilities of their state and those of its adversaries, may develop unrealistic expectations as to what the new military technology offers in terms of practicable capability gains. All things equal, the more novel a new technology and the less similar it is to existing technologies, the greater the challenges associated with transitioning in a

manner that takes maximum advantage of the strengths of a new technology while mitigating its limitations.

The aforementioned considerations concerning the deployment process are primarily about the qualitative capability gains, if any, that are accrued to a state through its deployment of a new military technology. There are, however, additional primarily quantitative considerations worth examining. While the throughput of the production process determines the pace and, on a per unit of time basis, the magnitude of effects on a military balance engendered by technological change, transitioning to a new technology comes with trade-offs that tend to lower a military's combat readiness for a time and thereby temporarily reduce its fielded military capabilities until the new technology can be employed by frontline units. Unless a state raises new/fresh forces to employ a new technology (which requires the injection of new/previously unallocated manpower), it will have to reallocate existing manpower and assign the new technology to existing forces/units. Although there is considerable variance in terms of the amount of training required to operate and maintain new a military technology, the personnel and units that are put through this transitioning process are likely to experience a temporary drop in readiness (to engage in combat operations). In a world of infinite resources including infinite manpower, a state could simply raise new units and later rotate these with the existing units so that, in time, its entire force structure could transition to the new technology. In the real world, of course, states rarely have such luxuries and instead make allocative choices that come with significant trade-offs. A state can, for example, accept a temporary decline in readiness for an entire unit and have complete units transition to the new technology one at a time.

Consider, for example, a situation in which State A has three tank brigades (with a triangular structure such that each brigade is composed of three battalions) and can transition to a new tank design (a new technology) one battalion at a time so that it can continue to employ eight of its nine tank battalions when required. If it is assumed that the transition time amounts to two months per battalion (for units and personnel who already operate a different tank design) and that the manufacturing process can deliver a battalion's worth of the new tank and associated equipment every two months (a rather optimistic assumption for a state that plans to only field nine tank battalions in total) so that there are no gaps between the retraining/transition for the nine tank battalions, it will take State A eighteen months to fully convert its three tank brigades to the new technology. State A may however view the operational risks resulting from a reduction in the readiness of one-ninth of its tank battalions too great particularly as most military units are not at 100% readiness in peacetime to begin with and having a single brigade lose access to a third of its tank battalions may therefore be unacceptable depending on the state's military-geographical context and its threat perceptions.

To deal with such a situation, State A could instead take a different approach and transition one company of a single battalion at a time (with each battalion having three companies). In so doing, and assuming the same transition capacity of one battalion every two months, each of the three tank brigades would lose only one-ninth of their tank companies at a time. While this can average out the temporary reductions in readiness across the three tank brigades such that each brigade would always have a maximum theoretical combat strength of 89% (as opposed to one in which the three brigades are at 66.6% strength), it results in a situation in which it would take eighteen months for each

brigade to complete the transition to the new technology. If State A were to transition one battalion from a single brigade every three months, then it could prioritize the conversion of one of its three brigades and thereby transition an entire brigade within nine months. What this brief and highly simplified analytical exercise highlights that is irrespective of which approach a state takes to deploy a new military technology, it encounters difficult trade-offs and a temporary decline in combat readiness – a temporary decline in its fielded military capabilities – before it can fully realize the capability gains nominally afforded by a new technology.

It is worth noting that the above example assumes that State A can access such an industrial capacity (whether in its own territory or in other countries) that can deliver a battalion's worth of the new tank design and associated equipment every two months, which is to say somewhere in the region of approximately 180-260 tanks per year, a figure that is exceedingly optimistic if not outright unrealistic for a country that only operates nine tank battalions, or approximately 280-400 tanks (i.e., the new tank design would only have a two-year production run assuming a constant rate of production and no export customers). Needless to say, the transition process will be longer if the production rate is lower, and, as such, it will take even longer – likely many years – for the capabilities gains nominally engendered by the adoption of the latest in tank technology to be realized. In the interim, our hypothetical state will deploy a hybrid force structure and will not therefore fully benefit from nominal gains in its fielded military capabilities even though the new tank design has already been developed, tested, and manufactured.

In the real world, the duration of the conversion/transition process can vary considerably across technologies but, all things equal, the more novel a technology, the

greater the challenges of transition are likely to be. If, for example, a state develops a new assault rifle that has a very similar design to its existing rifle and uses the same ammunition and magazines, then the training required to transition to the new technology is likely to be very limited. In such cases, frontline units may even be able to independently transition to the new technology without the need for specialized training personnel and facilities. This would not only eliminate a bottleneck in the deployment process, but it could also allow frontline units to remain at a high state of readiness while the transition takes place. In cases such as these, the deployment process does not greatly moderate the effects of technological change on the fielded military capabilities of states. In the real world, new technologies that offer substantial gains in capabilities unsurprisingly tend to be considerably different from the existing technologies that these are intended to replace. As such, the adoption of a new rifle that uses the same ammunition and magazines is unlikely to offer a significant gain in fielded military capabilities. There can be exceptions to this. For example, the new rifle may be very similar to the existing rifle, but the existing rifles may be decades old and increasingly unreliable as a result of years of mechanical wear. Alternatively, the new rifle may replace a broadly similar existing rifle that had major design defects that rendered it unreliable, thereby necessitating replacement. In such cases, a state may experience a gain in its fielded military capabilities, but it is important to recognize that this is not necessarily a function of any major performance growth in the new technology being procured. Stated differently, states can obtain a gain in capabilities by procuring additional quantities (i.e., newly produced specimens) of existing technology (because existing equipment is, once built and used, a depreciating asset) but, by so doing, the

nature of the capability gains on offer is distinguishable from those that can be obtained through the pursuit of a new design that reflects upon an expansion of the technological frontier and offers substantial gains in terms of performance growth.

Most new military technologies require some amount of training for operators and maintainers, and this results in a temporary drop in the readiness of the units receiving the new technology. In some cases, such as training in the use of a new radio, the transition process may take no more than a few days to a few weeks, at least for the operators. In other cases, however, the transition process can take many months if not several years. When, for example, a state adopts a new (jet) fighter aircraft, it may readily take six to eighteen months, if not longer, for a fighter squadron to (re)train its pilots, maintainers, and command personnel to such a level that the squadron can be characterized as having a full operational capability with the new technology (i.e., a state in which operators can fully utilize what the new technology offers). During this time, the combat readiness of the squadron transitioning to the new technology tends to drop close to zero and thereafter gradually increases toward 100%. The unit is therefore essentially irrelevant to the balance of fielded military capabilities during segments of the deployment process even as it is at the forefront of transitioning to a new technology that can ostensibly engender significant gains in the state's fielded military capabilities at a later date.

Because of the lengthy transition time commonplace when adopting more complex forms of military technology such as fighter aircraft, larger militaries with larger force structures tend to require multiple years, and in some cases the better part of a decade if not longer, to fully transition to the latest technology in a given area. At times, this drawn-out transitioning process can result in situations in which the final production

batch of an erstwhile “new” technology, being completed years after the first batch, may be increasingly outdated and the production line may shortly convert to the production of a successor design. Compounding matters for militaries, these dynamics entail the fielding and sustainment of hybrid force structures in which newer and older equipment must coexist and this tends to increase the logistical burden such that militaries tend to allocate more resources simply to sustain a hybrid force structure without benefitting from commensurate gains in their fielded capabilities. In a world of infinite resources, militaries would order quantity *z* of a new military technology and build and thereafter field it “all at once” but in the real world the scarcity of resources, allocative trade-offs, and other considerations such as readiness do not permit this.

The purpose of this highly simplified analytical exercise is to underscore that the existence of completed or manufactured specimens of a new technology does not inherently accrue a state gains in its *fielded* military capabilities and that the deployment process tends to further moderate the effects of technological change on fielded military capabilities through a variety of mechanisms including that concerning combat readiness.

3.4 A Dyadic Framework for Assessing the Pace and Magnitude of Shifts in Fielded Military Capabilities

The analytical framework developed in this chapter has so far focused on monadic mechanisms that set aside the possibility that other states can respond to the development of a new military technology by undertaking efforts to blunt if not negate its effects. Even without an adversary undertaking such efforts, there are multiple monadic factors and mechanisms that suggest that technological change is unsuited to the role of a prime

mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. Technological development tends to be a protracted and uncertain affair, particularly for more novel technologies that expand the technological frontier. The processes of testing, manufacturing, and deploying a new military technology are also time-consuming and far from straightforward. All things considered, the timeframe between the start of development and the start of fielding is often measured in years and this understates the duration for it excludes the timeframe over which prior, less effectual versions of a technology may have been developed as well as the time it takes for a state to produce, field, and integrate the new technology in such numbers that it constitutes a substantial portion of its frontline forces. For entirely new technologies for which many components also must be developed, the timeframes for development and successful integration of the new technology alongside the rest of a capability set tend to be longer still.

Although the monadic segment of this analytical framework highlights a variety of mechanisms through which the pace and magnitude of shifts in fielded military capabilities can be moderated, it is nevertheless possible that gradually accruing (absolute) gains in fielded capabilities eventually accumulate into capability shifts that are of greater significance for international politics. After all, in a world – that of a monadic analytical framework – in which there are no adversaries, military capabilities only exist in absolute terms and the primary obstacle to increasing capabilities are resources and time, and states will, in effect, be racing against themselves. In the real world, of course, many states are in adversarial relationships with one or more other states and any analysis of the implications of technological change for international politics is therefore incomplete without consideration of the effects of the competitive

interactions of states working to maintain their military capabilities in response to technological change, among other drivers of shifts in fielded military capabilities.

How can states avert declines in their fielded military capabilities? As indicated earlier, states have three general approaches through which to blunt or negate the effects of technological change. States can attempt to *adopt* the new technology. States can attempt to *counter* the new technology. States can attempt to *offset* the new technology. Before examining each of these approaches in turn, it is important to emphasize that these are not mutually exclusive even though the allocation of finite resources necessarily entails trade-offs and the attendant need for prioritization. It is also important to emphasize that states do not necessarily lock themselves into these approaches and that real-world military-technological competitions tend to have multiple dynamic pathways through which states can blunt if not negate the effects of technological change on their fielded military capabilities over time. The remainder of this chapter will develop a dyadic framework for understanding the mechanisms through which technological change can affect the fielded military capabilities of states.

3.4.1 Option A: States can adopt a new technology.

Although much of international relations scholarship does not concern itself with the development of new technologies and instead focuses on the effects, broadly defined, of technological change, there is nevertheless a widespread presumption that a new technology will eventually diffuse. This diffusionist presumption is so strong that the military innovations literature, which explicitly concerns itself with the patterns of adoption and diffusion of new military technologies and doctrines for the employment

thereof, discusses, for example, German tanks and tank doctrine and compares these with British, French, American, and Soviet tanks and tank doctrine but make little if any mention of the possibility – and the historical reality – that these states did not necessarily need tanks of their own to blunt if not negate the effects of another state's introduction of tanks or otherwise enjoy gains in their fielded military capabilities by adopting other types of armored combat vehicle. Stated differently, international relations scholarship tends to view the concurrent adoption of a new technology as a dynamic indistinct from countering – it tends to view a tank as a counter, if not the only counter, to another tank and, in effect, tends to view the adoption of a new technology as a means of countering its effects.

Although there is a case to be made for countermeasures (countering will be examined in the next section), it is imprudent to respond to the diffusionist tendencies of international relations scholarship with a rejection of the possibility that adoption may (also) offer states a means to blunt if not negate the effects of technological change on their fielded military capabilities. By adopting a new technology that State B is first to introduce, State A can obtain the same (monadic and absolute) capability gains as State B, and this can help blunt if not negate the net effects of the technological advance on the State A-State B military balance. If State A was already pursuing this area of technology in some way, shape, or form, then State A's adoption of the new technology is best characterized as an act of convergence. If State A was not pursuing the new technology before State B introduced it, then it is best to characterize State A as emulating State B. This distinction is subtle but important for it reflects the reality that states do not always have an unambiguous lead in introducing a new technology and that their adversaries

may not be far behind given that scientists, engineers, and military officers largely encounter the same scientific and engineering opportunities and constraints. As a result, the capability gains engendered by introducing a new technology may be non-zero, but these gains may not be restricted to and fully monopolized by one state and the net effects of technological change on a military balance may be very different from the absolute and monadic effects that may exist in a world without adversaries.

3.4.2 Option B: States can counter a new technology.

Adopting a new technology may offer a means of blunting if not negating the effects of technological change on a military balance but it may not be the most effective and/or most cost-effective approach that a state can pursue. Consider, for example, a situation in which State A responds to State B's development and fielding of tanks by adopting tanks of its own. If equipped with suitable armament, State A's tanks could damage if not destroy State B's tanks in combat and the capability gains accrued to State B through its development and fielding of this new technology could therefore be moderated. Having adopted tanks of its own, State A could also employ the new technology in other roles, including more general offensive and/or defensive roles that do not necessarily involve using tanks to target State B's tanks. In other words, State A could adopt tanks and in so doing blunt if not negate the effects of State B's introduction of a new military technology on the State A-State B military balance whether by directly using tanks against other tanks or by benefitting from capability gains that can offset State B's capability gains by using tanks against other, non-tank, area of State B's capability set (offsetting will be examined in the next section).

In a world of unlimited resources and no trade-offs in resource allocation (or in the design of military technologies, more of which will follow), states would encounter incentives to pursue the "best" technology available for each and every possible role. In the real world, states do not have such luxuries and must instead prudently allocate their finite resources. While states can adopt technology *x* and plan to use it (in a symmetrical manner) against an adversary's version of the same technology, states encounter incentives to develop more effective and/or more cost-effective means of countering a new technology. Returning to the example of tanks, State A could deploy such a number of tanks that it would have enough tanks to blunt if not negate State B's tank force *and* (still) employ its tanks in other roles, whether offensive or defensive. Tanks are, however, expensive to build and operate, and State A may not be able to afford so many tanks for all possible roles and would still encounter allocative trade-offs (concerning technologies *y*, *z*, etc.) even if it had the resources to do so. And if State A does not have the resources to develop tanks of its own let alone build and deploy tanks in militarily significant numbers, then exclusively pursuing the adoption of tanks would not be a particularly prudent approach through which to blunt or negate the effects of State B's deployment of the new military technology on State A's fielded military capabilities.

Rather than symmetrically adopting tanks as its only means of blunting the effects of the latest in tank technology fielded by State B, State A could pursue more specialized capabilities intended to target or counter State B's tanks. State A could, for example, develop human-portable anti-tank armaments that can be more affordably produced in large numbers and deploy these across its frontline units, thereby increasing both the quantity and the density of anti-tank capabilities that State B's tanks would encounter.

State A could still pursue tank technology (i.e., adoption) but separate the ordnance (armament) of a tank (assuming that this is suitable for use against State B's tanks) from the chassis and instead deploy this armament on a towable mounting (i.e., develop anti-tank artillery) or even employ such ordnance in a (stationary) bunker for use against State B's tanks. State A could also decide not to develop and build a single type of armored combat vehicle (i.e., a tank) and instead develop and build some number of different and more specialized self-propelled anti-tank vehicles that typically trade speed and some level of mobility for more powerful armament and perhaps greater levels of (frontal) armor protection to be more effective in an anti-tank role at the cost of being less effective in other roles for which tanks may be more suited (typically offensive roles involving fire and maneuver for which self-propelled anti-tank guns tend not to be well suited). Provided that it has adequate resources, State A could of course pursue all of these (and other) approaches in addition to adopting tanks of its own (as could State B and other states).

What these examples of how a state can respond to the introduction of a new technology such as tanks highlight is that states tend to have ample opportunities to develop and field *technological countermeasures* through which to blunt if not negate the capability shifts otherwise engendered on a military balance by technological change. These technological countermeasures to a new technology can vary considerably and may incorporate either components of the new technology or the new technology in its entirety (albeit in ways that distinguish it from "mere" adoption). For example, a dedicated (gun-equipped) tracked anti-tank vehicle is essentially a derivative of tank technology optimized for a specific (i.e., anti-tank) role while a towed anti-tank gun only

draws upon the armament of a tank and not its chassis and propulsion system. Here, we encounter the possibility of overlap between the approaches of adopting a new technology and countering it, but it is important to recognize that what distinguishes the approach of countering from that of adoption is that a state does not merely pursue the new technology as an end in itself or intend to obtain the same prospective capability gains that the new technology is accrued to another state. The option of countering is, in other words, distinguished by the deliberate development of capabilities to target a new technology including, but not limited to, countermeasures that may have little to no purpose other than serving in the anti-new technology *x* role. Even if a technological countermeasure can serve multiple purposes, there may be little to no overlap between the approaches of adoption and countering. For example, a shoulder-fired anti-tank weapon can be used against other types of targets, but it does not share the armament, armor protection, propulsion system, or any other significant component of tank technology and cannot as such be characterized as a manifestation of the adoption approach in response to the introduction of the latest in tank technology by another state. Stated differently, as far as the introduction of tank technology is the point of reference, shoulder-fired anti-tank weapons are exclusively countermeasures to tanks (although shoulder-fired anti-tank weapons can also effectively augment the firepower of dismounted infantry against structures and other infantry).

Although this analytical framework posits that states can pursue (technological) countermeasures through which to blunt if not negate the effects of a new technology on their fielded military capabilities, it is not necessarily obvious why such countermeasures are possible in the first place or why the feasibility of countering should be taken as a

given as implied thus far in this text. Simply put, why is a new technology likely to be vulnerable to countermeasures in general and technological countermeasures in particular?

The designers of all military technologies encounter physical constraints – which are essentially immutable – and must navigate these with the technologies available to them (within the contemporaneous technological frontier) and the result of this interaction is that all human technologies are necessarily compromised and reflect design trade-offs. The designers of a new tank design, for example, may want the highest level of armor protection possible but are nevertheless limited by the most capable propulsion system and drivetrain available to them. Even if these engineering constraints emanating from the state of technology (i.e., the technological frontier) pose no obstacle, the designers must consider the weight of the vehicle if they place any value on its mobility. Vehicular armor is heavy, the maximum power output of an engine is finite for a given engine design, and increasing the weight of a tank without a commensurate increase in engine power output results in a lower power-to-weight ratio which is manifested in terms of a lower top speed, slower acceleration, and so forth. Designers must also consider both natural and anthropogenic physical constraints. To effectively operate offroad including on soft ground, tanks must have low ground pressure, and this puts practical constraints on the maximum weight of a tank which in turn means that there is an upper limit on the amount (i.e., weight) of armor – and thus the level of armor protection – possible with a given tank design. Designers could, of course, create a dimensionally larger vehicle to reduce the ground pressure to acceptable levels while incorporating the desired higher level of armor protection. Doing so would however run

into anthropogenic (physical) constraints in that the resulting tank design could be too heavy and/or too large to traverse existing bridges and make use of existing transportation infrastructure more generally. The implication of these competing and incompatible requirements is that all technologies reflect design trade-offs and are therefore compromised in one or more respects.

Given that the designers of a new technology such as new tank tend to encounter competing and incompatible design priorities, they must deliberately optimize the design on certain performance parameters and accept the associated trade-offs and, as such, the attendant vulnerabilities. Designers and military personnel looking for blunt if not negate the effects of a new technology such as a new tank design have every incentive to identify the physical and engineering/technological constraints that tank designers encounter and use this information to make informed judgments about the likely trade-offs in a given tank design and thereby identify areas of weakness against which to target the development of countermeasures. Given the limitations on vehicular weight and dimensions, for example, tank designers have to prioritize armor protection in certain parts of the vehicle at the expense of others. For practical reasons, essentially every tank design emphasizes frontal protection under the assumption that this is the direction from which a tank is most likely to encounter enemy fire (the alternative to optimization would be to have all sides of the vehicle equally vulnerable to enemy fire). The result of this optimization and the accompanying trade-off is that essentially every tank ever designed has tended to be resilient against contemporary anti-tank capabilities targeting the front of the vehicle while being more vulnerable on the sides as well as in the rear, on the roof, and the bottom of the vehicle. Consequently, states have developed countermeasures

intended to primarily target these specific areas of a tank. Infantry equipped with shoulder-fired anti-tank armaments, for example, are typically instructed to prioritize attacking the side armor when possible since it tends to be less well protected relative to the frontal armor. To target the relatively weak armor protection on the underside of a tank, militaries have developed a wide array of anti-tank mines. To overcome increasingly effective frontal and side armor in the latest tank designs, militaries have developed top-attack munitions that target the less well-armored roofs of tanks and other armored vehicles. As successive generations of tank designs introduced greater levels of armor protection, the size (i.e., diameter) of the warheads of these anti-tank armaments, which are typically of the shaped charge variety, has increased in tandem. What these examples indicate is that the interaction between physical constraints and the technological frontier beget design trade-offs and design trade-offs beget potential vulnerabilities to countermeasures.

This discussion of countermeasures and the countering approach has so far focused primarily on technological countermeasures. It is important to recognize, however, that states need not necessarily develop and field any new military technologies in order to avert sizeable negative shifts in their fielded military capabilities emanating from technological change. States can modify the tactics employed by their military personnel, reform their organizational practices, and adjust their military doctrine. These approaches are not mutually exclusive, and it is possible for a state to pursue some combination of tactical-organizational-doctrinal adaptations (countermeasures) alongside technological countermeasures (as well as adopting and/or offsetting a new technology).

Most if not all technological countermeasures are accompanied by new or revised tactics, organizational structures, and doctrines regardless of whether the technological countermeasure is a new technology or an adaptation of existing technology. For example, the advent of progressively more capable shoulder-fired anti-tank munitions and guided anti-tank munitions led states to revise infantry tactics, disestablish anti-tank gun units, establish new anti-tank units equipped with shoulder-fired anti-tank weapons as well as larger, longer-ranged, and less mobile versions of the same technology, and, not least, revise their doctrinal approaches concerning both defensive and offensive operations in a manner that reflected upon the proliferation of anti-tank capabilities across the battlefield. For example, rather than simply setting up anti-tank guns to lie in ambush, advances in anti-tank capabilities allowed light infantry units and lightly armored and unarmored vehicles equipped with anti-tank recoilless guns and anti-tank missiles to more actively hunt for tanks on the battlefield and this is an example of a technological development – a technological countermeasure – allowing for new tactical-organizational-doctrinal adaptations to countering tanks on the battlefield.

3.4.3 Option C: States can seek to offset the new technology.

States attempting to blunt the effects of technological change on their fielded military capabilities can adopt said technology and/or seek ways to counter it. There is, however, an often underappreciated third approach that states can elect to pursue with or without adopting and/or countering the new technology – attempting to offset the *effects engendered* by new technology. The offsetting approach encompasses both attempts to supplant a new technology – undertaking efforts to make it less relevant to the military

balance – and pursuing capability gains in other areas of a military balance to, on balance, blunt if not negate the *net capability gains* otherwise accrued to an adversary through its deployment of new technology. Offsetting can entail technological and/or organizational-doctrinal innovation but can also entail further resource allocation toward existing capability sets and can also be an unexpected product of unwitting actions. Stated differently, offsetting is a residual approach category that excludes doing nothing but encompasses a wide range of options distinct from adoption and countering through which states can work to maintain their aggregate military standing even if they are relatively disadvantaged in one or more subareas of a military balance as a result of technological change.

Recognizing offsetting as a distinct approach that states can respond to technological change is important because military balances tend to be complex and multifaceted, and states rarely allocate their resources in a truly symmetrical manner. Military balances are, as such, rarely restricted to a single area of technology and it is therefore problematic to assume that states must only adopt or counter a new technology. To appreciate the benefits of viewing offsetting as a distinct approach, it is useful to continue the example of how states can respond to advances in tank technology.

When State B introduces a new technology such as tanks, it benefits from gains in specific areas of its fielded military capabilities. With tanks, State B may be better positioned to undertake offensives against State A or – and this is not mutually exclusive – more effectively defend against offensives undertaken by other states against it. With its introduction of tanks, State B does not benefit from gains in other areas of its military capabilities such as its ability to employ surface shipping and deny its adversaries from

doing the same or, for example, benefit from an increase in its air defense capabilities. State A could attempt to avert a net decline in either its offensive and/or defensive capabilities relative to State B by adopting and/or countering tank technology but in doing so State A would be engaging in a (new) area of military-technological competition initiated by State B and, at least at the outset, would do so on State B's terms. This may or may not be a prudent approach for State A given its comparative advantages, broadly construed, its resource base, industrial capacity, demography, geography, preferred military strategy, or political and/or diplomatic considerations, among other things.

It is possible, for example, that State A does not think it can afford to allocate adequate resources toward adopting tanks and/or countering tanks or that the opportunity costs of doing so could not be justified. State A may therefore elect not to maximally engage in this new area of military-technological competition and instead seek to blunt if not negate the *net* effects of State B's introduction of tanks on the State A-State B military balance by opening, or doubling down on, one or more new areas of military-technological competition or allocating its finite resources in other existing areas of military-technological competition. By doing so, State A would be attempting to *offset* the net capability gains accrued to State B by its introduction of new technology such as tanks. For example, State A may decide to invest in improving its capacity to undertake airstrikes in State B's territory. By so doing, State A may incidentally benefit from an increased capacity to target State B's tanks from the air, but State A may be focused on working to offset its relative decline in ground combat capabilities brought about by State B's introduction of the latest in tank technology by pressuring State B elsewhere. By so doing, State B would face incentives to reallocate its finite resources toward bolstering its

air defense and, as such, may have to cut back on its procurement of new tanks (or failing that, cut back and/or forgo investment in some other area).

To be clear, there is no guarantee that offsetting will have the desired effects on an aggregate military balance, but the point nevertheless stands that is analytically unproductive to restrict the menu of options available to states in the face of an adversary's adoption of a new military technology to a binary of adopting the new technology and countering the new technology. While these two options may characterize the menu of state options in a highly disaggregated assessment of a subset of the aggregate military balance, a more holistic, that is aggregated, appraisal of a military balance is invariably multifaceted such that, in simplified terms, gains in, for example, ground combat capabilities do not necessarily translate into gains in naval or air combat capabilities. As such, adversaries may pursue heterogenous military strategies, technological priorities, and force structures as part of multi-faceted and dynamic efforts at offsetting one another's capabilities, particularly in areas of the military balance in which their adversary may have a comparative advantage. By considering the possibility of offsetting approaches as a means of blunting if not negating the effects of technological change on an *aggregate* military balance, we benefit from a more holistic conception of military balances that does not artificially reduce military balances to highly disaggregated areas of competition.

3.5 Analytical Framework Conclusions

The analytical framework developed in this chapter investigates the mechanisms through which advances in military technology translate into shifts in the fielded military

capabilities of states and whether these shifts are sizeable and abrupt. Shifts in fielded military capabilities are not abstractions and require concrete indicators and mechanisms. The analytical framework developed here covers both monadic and dyadic mechanisms through which technological change can affect the fielded military capabilities of states from the developmental process to the deployment process.

3.5.1 Is there a direct mechanism through which technological change can engender shifts in fielded military capabilities?

The analytical framework developed in this chapter identifies a variety of mechanisms through which advances in military technology can drive shifts in the fielded military capabilities of states but the very act of disaggregating the monadic mechanisms – development, testing, production, deployment, and post-deployment integration – highlights how tenuous and contingent the effects of technological change can be. Even in a purely monadic framework in which no adversaries are attempting to blunt if not negate the effects of technological change on a military balance, these mechanisms constitute bottlenecks regulating the throughout – both the pace and magnitude – of the effects of technological change on a military balance. In practical terms, it typically takes multiple years for a state to develop, test, manufacture, and deploy a new technology in such numbers as to have a significant effect on a military balance in a state dyad. The dyadic mechanisms examined in this chapter highlight how technological change has no intrinsic effects on a military balance as the effects are coproduced by the actions taken and forgone by an adversary. All things equal, the more time that an adversary state has to respond, the greater the opportunities to adopt, counter, and offset the new technology

and thereby blunt if not negate the effects of technological change on its fielded military capabilities. In practical terms, countries in a responding state's position tend to have considerable opportunities to undertake efforts to blunt if not negate the effects resulting from another state's introduction of a new technology.

Given this, the answer to the question of whether there is a direct mechanism through which technological change engenders shifts in fielded military capabilities is a decidedly qualified yes. That is, there are plausible pathways for effects, but these are highly contingent such that is it problematic to assume that technological change tends to engender militarily significant shifts in fielded capabilities that are of consequence for international politics. If anything, both the monadic and dyadic frameworks suggest that, as with other candidates for the role of a prime mover engendering sizeable and abrupt shifts in fielded capabilities, effects tend to be lagged and indeterminate as whatever the qualities of a new technology and the envisaged capability gains that it offers an adopting state, the practical real-world effects are neither inherent nor automatic and are instead coproduced by the actions taken and forgone by adversaries in complex multi-faceted military competitions in which the importance of military technology is often overshadowed by the realities of resource allocation, among other important considerations.

3.5.2 Does change in this area – technological change – lend itself to driving shifts in fielded military capabilities that are both sizeable and abrupt?

Even if there are direct mechanisms through which technological change can drive shifts in fielded military capabilities, the significance of a non-zero shift in fielded capabilities

for international politics is not intrinsically endowed. It is, rather, a function of the pace and magnitude at which shifts in fielded military capabilities unfold. With the effects coproduced by the states in a dyad, both the pace and magnitude of change are important, and a suitable candidate for the role of a prime mover must have direct mechanisms to affect fielded military capabilities but also lend itself to engendering shifts in fielded military capabilities that are both sizeable and abrupt.

Does technological change lend itself to driving sizeable and abrupt shifts in fielded military capabilities? The monadic mechanisms identified in this framework highlight how the effects of technological change on fielded military capabilities are moderated by the processes of development, testing, manufacturing, and deployment. Each of these processes constitutes a bottleneck and tends to be prolonged. The developmental process, for example, is typically a multi-year-long affair and tends to be even longer once an area of technological development in a given area is viewed longitudinally. The discussion of the testing process highlights the risks associated with forgoing thorough testing and trials before committing a new technology to production. The discussion of the manufacturing and fielding processes highlights the bottlenecks in the effects of technological change on fielded military capabilities that states encounter even in cases in which a new technology is "ready" for frontline service. Finally, the discussion of the deployment process highlights how the mere existence of completed specimens of a new technology ready for use by frontline military units does not automatically engender gains in the fielded military capabilities of states absent considerable time and effort.

While these monadic mechanisms highlight the bottlenecks in translating technological change into fielded military capabilities in a world in which there are no adversaries, it is the dyadic segment of the analytical framework developed in this chapter that captures the degree to which technological change tends not to engender shifts in capabilities that are both sizeable and abrupt. In the hypothetical world captured in the monadic segment of the analytical framework, the only obstacles to maximal gains in fielded capabilities being accrued to a state are resources and time. In the real world, other states have incentives to adopt, counter, and offset a new technology in order to maintain their military capabilities. As a result, the effects of technological changes introduced by State A on the State A-State B military balance are only as sizeable and abrupt as State B, to say nothing of other states outside this highly stylized two-actor model, allows. The effects are, in other words, coproduced by the interaction of both monadic and dyadic mechanisms and this offers considerable scope for moderation such that the effects engendered on fielded military capabilities may be neither sizeable nor abrupt.

The analytical framework developed in this chapter suggests that technological change is not a suitable candidate for the role of a prime mover engendering shifts in the fielded military capabilities of states that are both abrupt and sizeable. This, however, is a purely analytical claim. Does the empirical record support this characterization or does it indicate otherwise? In Part Two and Part Three of this dissertation, which encompass chapters respectively examining advances in submarine and uncrewed aircraft technology, I turn to the historical record to determine whether this analytical framework can be empirically substantiated.

Part Two: Advances in Submarine Technology ca. 1800-1945

Chapter 4. Submarine Technology: Setting Up the Analysis

This chapter examines whether the experience with advances in submarine technology in the ca. 1800-1945 timeframe provides support for the candidacy of technological change as a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. Before proceeding to longitudinally examine the history of submarine technology and the effects of advances therein on the fielded military capabilities of states, we must make two interrelated analytical determinations. The first concerns the establishment of a baseline through which to discern the effects of technological change and the second concerns the temporal scope of this analysis.

Given the aim of discerning whether there is empirical support for the notion that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states, this chapter primarily examines the degree to which submarines *supplanted* surface warships in the capability sets fielded by militaries. By supplanting, I refer to the process by which technological change allows a military to gain from replacing an existing capability set wholesale. There are multiple examples of this dynamic of (eventual) supplanting in military history. Horse cavalry, for example, was wholly supplanted by armored combat vehicles and unarmored and/or unarmed vehicles of various types, including tanks, albeit after several decades of gradual technological development during which such self-propelled vehicles of various types could productively coexist with horse-mounted units (including mounted infantry) in force structures. Similarly, lighter-than-air aircraft, which is to say balloons, blimps, and rigid airships, were wholly supplanted in combat roles (and, in time, were supplanted in essentially every conceivable role) by fixed- and rotary-

wing heavier-than-air aircraft. As these examples highlight, the degree to which a new military technology supplants existing technologies and capability sets constitutes an important indicator of the effects of technological change on the fielded military capabilities of states. Although new military technologies can offer militaries the ability to undertake important new missions, the implications for fielded military capabilities will necessarily be limited unless and until the new technology can undertake existing missions (or if new missions emerge that supplant the existing missions or otherwise make existing missions less relevant). If a new technology does not supplant existing technologies, it will likely coexist in a hybrid force structure alongside *existing* technologies, and this will inherently moderate the overall implications of technological change on the fielded military capabilities of states.[1]

It is important to emphasize that this chapter's examination of whether advances in submarine technology allowed submarines, the "new" military technology, to supplant surface warships, the existing military technology, is not an arbitrary standard crafted for the present purposes but one that reflects upon a century or so of writings heralding an ostensible consequential demise of surface warships as a result of technological change. By the early twentieth century, the increasing prominence of (naval) mines, torpedoes (irrespective of launch platform), and submarines led to intense debates about the viability of large surface warships including battleships. During the Interwar Period and throughout the Cold War, strikingly similar debates transpired although the technological focus increasingly turned toward the threat posed to surface warships by crewed fixed-

[1] While new military technologies tend to coexist alongside existing technologies until the former can be produced and fielded in such numbers as to replace the latter, it bears emphasis that the dynamic being highlighted here is not driven by the scarcity of resources but the unsuitability of a new technology to wholly replace an existing technology.

wing military aircraft (including torpedo bombers), whether based on land or an aircraft carrier, and, in the Cold War period, the threat posed by anti-ship missiles of various types that supplanted torpedoes in the aerial targeting of surface ships.[2] Attesting to the longevity of this debate as to whether technological change is driving the supplanting of surface warships (i.e., "making surface warships obsolete"), more recent writings focus on a wide range of so-called "anti-access capabilities" that encompass submarines and anti-ship missiles of various types including anti-ship ballistic missiles.[3] For better or for worse, writings on the implications of technological change in the naval realm have long been characterized by debates as to whether technological change in one area or another allows for a new technology to supplant an existing technology – surface warships – and the experience with submarine technology in the late nineteenth and early twentieth centuries provides one of the earliest examples of such debates.

While the preceding paragraphs establish a baseline through which to discern the effects of technological change on the fielded military capabilities of states, we also need to determine the temporal scope of analysis. Many writings on the implications of advances in submarine technology begin with the First World War, which was the first conflict in which self-propelled submarines were used in combat (human-powered submarines were, however, employed in the American Civil War).[4] Whatever the effects of submarine technology on the military capabilities of participants in the First World

[2] Norman Friedman, *Fighters Over the Fleet: Naval Air Defense from Biplanes to the Cold War* (Annapolis, MD: Naval Institute Press, 2016)
[3] See, for example, the debates covered in Stephen Biddle and Ivan Oelrich, "Future Warfare in the Western Pacific: Chinese Anti-access/Area Denial, U.S. AirSea Battle, and Command of the Commons in East Asia," *International Security*, Vol. 41, No. 1 (2016): 7-48.
[4] "H.L. Hunley Wreck (1864)" Naval History and Heritage Command, no date, https://www.history.navy.mil/research/underwater-archaeology/sites-and-projects/ship-wrecksites/hl-hunley.html

War, it is analytically problematic to begin an analysis of this conflict as submarine technology did not emerge in a vacuum in the lead-up to 1914. We must, as such, conduct a longitudinal analysis to discern the implications of advances in submarine technology over time leading up to the First World War before examining the implications of the use of submarines in that conflict let alone the implications of advances in submarine technology in later periods. Whereas no tanks, that is, tracked armored combat vehicles, existed before the First World War, and whereas powered fixed-wing aircraft had only existed for barely a decade, submarine technology existed for over one hundred years before this "new" technology was first used in combat and by 1914 submarines were in military service for the better part of twenty years. There is, as such, no sound reason to truncate the analysis such that submarine technology is treated as if it emerged out of the blue in 1914, especially since doing so creates the inaccurate impression that the "sudden" emergence of submarines affected the (wartime) fielded military capabilities of states to some extent or another.

To accurately discern the pace and magnitude of shifts in fielded military capabilities engendered by advances in submarine technology in the lead-up to the First World War and later timeframes, we must first examine the origins of submarine technology. If the developmental history of submarine technology is found to be characterizable by rapid and consequential technological development, then this area of military technology may provide empirical support for the notion that technological change engenders sizeable and abrupt shifts in the fielded military capabilities of states. If, however, the developmental history of submarine technology in the lead-up to the First World War is characterized by gradual and largely inconsequential technological

developments, then this undermines empirical support for the notion that technological change engenders sizeable and abrupt shifts in the fielded military capabilities of states irrespective of the implications of submarine technology on the First World War and later history (which will be examined later in this Part Two of this dissertation).

Justified though a longitudinal approach toward analysis may be, it does not establish a start date for analysis. A logical starting point for tracing the implications of advances in submarine technology over time would be with the "first submarine." The issue, however, is that there are several potential starting points – there are several candidates for the "first submarine – each separated by several decades. To discern a suitable starting point, I will first define what a submarine is for the present purposes. A submarine is, in its simplest form, a crewed maritime vessel capable of (re)submergence and both surfaced and underwater locomotion. A submarine is distinguished from a similar and closely related technology, the (self-propelled) torpedo, in two respects: (A) the submarine, being crewed, is not an expendable munition but a platform carrying a human crew and equipment in furtherance of one or more missions (ideally more than once – i.e., a reusable vessel); (B) a submarine is designed to operate laterally on the surface and subsurface as required.[5]

For the purposes of this dissertation, I propose setting 1800 as the year in which a military submarine became a *possibility*. In that year, one of the earliest well-attested submarine designs, the American Robert Fulton's *Nautilus*, underwent tests (off the coast of France), served as a proof of concept, and defined the technological frontier. As I will explain in the following chapter, *Nautilus* and submarine designs of its generation had

[5] This second criterion also distinguishes a submarine from equipment like a diving bell that allows humans to submerge but which do not provide a means of lateral mobility.

negligible effects on the fielded military capabilities of states and more practical submarines would not emerge for many decades. Nevertheless, the advent of the *Nautilus* marks perhaps the earliest possible time in which militaries *could have* pursued submarines if they were so inclined, and, on that basis, this is a productive starting point for establishing – anchoring – a baseline and discerning the effects of technological change over time.

The remainder of Part Two of this dissertation is organized as follows. Chapter Five examines whether advances in submarine technology in the ca. 1800-1900 timeframe engendered sizeable and abrupt shifts in the fielded military capabilities states. By the start of the twentieth century, self-propelled submarines were entering frontline service in a growing number of militaries and Chapter Six extends the analysis into the ca. 1900-1914 timeframe. Having established the pre-war origins of the submarines employed in the First World War, Chapter Seven examines the military effects of the employment of submarines during that conflict. Chapter Eight extends the longitudinal analysis into the Interwar Period and Chapter Nine concludes Part Two of this dissertation with an examination of the implications of both the military effects of submarines in the Second World War as well as the implications of late-war advances in submarine technology. Before proceeding, it is productive to explicate that Part Two of this dissertation does not offer an operational history of submarine operations in the ca. 1800-1945 timeframe or examine the political and/or diplomatic effects of advances in submarine technology in times of peace or war. The following empirical chapters exclusively examine advances in submarine technology as a case study through which to probe the plausibility of the claims made in Chapter Three concerning the suitability of

technological change as a candidate prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

Chapter 5: Advances in Submarine Technology ca. 1800-1900

Having established an analytical baseline, a temporal starting point, and a definition of what a submarine is, we can now examine whether advances in submarine technology in the ca. 1800-1900 timeframe engendered sizeable and abrupt shifts in the fielded military capabilities of states. If technological change is a suitable candidate for the role of a prime mover engendering shifts in the fielded military capabilities of states that are both sizeable and abrupt, then we should find evidence of this in the important area of submarine technology.

Although the first half of the nineteenth century tends to receive little attention in many writings on submarine technology, it is problematic to select cases or select the temporal scope of a study on the basis of the outcome of interest. As I will explain in the following pages, decades of advances in submarine technology amounted to negligible if not non-existent effects on the fielded capabilities of states largely as a result of the enduring immaturity of submarine technology following the periodic building and testing of functional submarine designs. After a brief discussion of the origins of submarine technology in the early nineteenth century, I will focus more closely on the period ca. 1860-1870 in which, among other things, both self-propelled submarines and self-propelled torpedoes were developed, as well as the final decade of the nineteenth century and thereby examine the military implications of a late-century breakthrough in submarine technology that was many decades in the making. In so doing, I will highlight how long it took for submarine technology to reach ca. 1900 levels and examine the cumulative effects of technological advances on the fielded military capabilities of states by the start of the twentieth century, which is approximately the timeframe at which

speculation about the implications of advances in submarine technology began in the lead up to the First World War. Although it is tempting to be dismissive toward this century-long period (ca. 1800-1900) of negligible if not non-existent effects on fielded military capabilities, doing so misses the point that advances in military technology tend not to engender sizeable *and abrupt* shifts in the fielded military capabilities of states and this nineteenth-century experience is, as such, far from atypical.[6]

As noted earlier, there are multiple candidates for the "first submarine." In many respects, the first well-attested "submarine" was the *Turtle*, an oversized-barrel-like human-powered submarine carrying a crew of just one that was launched in 1775.[7] Although *Turtle* was employed in the American Revolutionary War, it was an exceptionally primitive design, and the American Robert Fulton's *Nautilus*, which was tested in 1800 in France (with the French navy being the intended customer), constitutes a more productive starting point as this was an armed submarine intended to undertake combat roles. The *Nautilus* was a human-powered submarine operated by a crew of three that hand-cranked a propeller so that the submarine, after somehow finding a target vessel and accurately navigating toward said target vessel, could deploy a primitive early naval mine that would be dragged toward a target ship.[8] On paper, the development and testing of the *Nautilus* offered militaries the option of undertaking underwater attacks against

---

[6] For a notable exception to the tendency to overlook earlier timeframes, consider Karl Lautenschlager, "Technology and the Evolution of Naval Warfare," *International Security* 8, No. 2 (1983): 3-51.; Karl Lautenschlager, "The Submarine in Naval Warfare, 1901-2001," *International Security* 11, No. 3 (1986): 94-140. See also Bernard Brodie, *Sea Power in the Machine Age* (Princeton, NJ: Princeton University Press, 1941)

[7] "The Submarine Turtle: Naval Documents of the Revolutionary War," Naval History and Heritage Command, November 2, 2017, https://www.history.navy.mil/research/library/online-reading-room/title-list-alphabetically/s/submarine-turtle-naval-documents.html

[8] "NH 2058 Drawing of Robert Fulton's NAUTILUS, 1798," Naval History and Heritage Command, no date, https://www.history.navy.mil/content/history/nhhc/our-collections/photography/numerical-list-of-images/nhhc-series/nh-series/NH-02000/NH-2058.html

surface ships at a time when navies, most notably the British navy, could regularly enforce crippling close blockades of ports.

The French navy, for which Fulton first built and tested the *Nautilus*, did not adopt this or any other submarine design even though the proof of concept was generally successful. Desperate though the French navy may have been to find means of defeating the blockade imposed by the British navy and innovative though the design of the *Nautilus* may have been, this pioneering vessel was nevertheless little more than a small submersible in which the human operators cranked a screw for underwater propulsion and otherwise relied upon a sail for surface transit. As a result, the speed, range, and endurance of this and other human-powered submarines were exceedingly limited. More generally, *Nautilus*' design, construction, and the state of metallurgy in this timeframe did not allow for submerged operations beyond several meters beneath the surface and the vessel was unsuited for operations in rough seas or much further than the shallows in the vicinity of a harbor.

Even if such material deficiencies could be addressed and leaving aside the intrinsic limitations imposed by the reliance on human physical exertion for submerged locomotion (and surface operation with a small sail unfurled), there was no practical armament for submarines in this timeframe. The *Nautilus* was intended to tow what amounted to a primitive early naval mine toward its intended target before releasing the tow rope and, ideally, escaping the blast. There was no means of underwater sensing to detect a potential target and there was no means of communicating with the submarine so as to provide the location of potential target ships. The *Nautilus* was, as such, limited to laying in ambush in the hope that a passing ship would be close – and slow – enough to

allow the human-powered vessel to reach its hull (to deploy the mine toward the target ship) or, alternatively, attacking enemy ships navigating a harbor or estuary, or, similarly, attacking ships lying at anchor close offshore. Given the limited effectiveness of a single *Nautilus*-type submarine and the limited range and endurance of any human-powered submarine, a sizeable fleet of such submarines would be required to credibly undertake coastal defense missions and enemy ships could for the most part safely operate by not casting an anchor close to an enemy-held port.

The limitations of human-powered submarines like *Nautilus* were glaring and militaries did not give serious consideration to the various design proposals that they periodically received across the first half of the nineteenth century even as the advent of marine steam propulsion and the subsequent armoring of warships progressively increased the perceived need for a means of defeating the latest large surface warship without a comparable large surface warship of one's own. Nevertheless, what Fulton (and others) demonstrated was that a military submarine *was possible* even if it was impractical and, all things considered, not worth pursuing. Stated differently, militaries *could have* pursued submarine technology from ca. 1800 onward (if not earlier) but did not, and, importantly, this was not because militaries were inherently opposed to adopting new technologies. After all, from the 1820s onwards, militaries invested heavily in new types of gun-type armament for warships and coastal defenses and, around the same time, adopted marine steam propulsion. By the 1850s, leading militaries invested heavily in making the most of advances in marine steam propulsion and blunting the effects of advances in naval guns by building self-propelled armored warships of various types. A lack of resources or a lack of interest in adopting or countering the latest advances in

naval technology cannot, therefore, account for the non-adoption of submarines in this timeframe. All things considered, the immaturity and impracticality of submarine technology, particularly human-powered submarines, from ca. 1800 through 1860 meant that submarine technology would likely have had minimal effects on the fielded military capabilities of states even if submarines had been built in sizeable numbers and fielded in frontline units (as opposed to the periodic building and testing of prototypes).

By the mid-1860s, militaries developed a renewed interest in submarine technology. Although there was a particularly notable development in submarine technology during the early 1860s, which will be examined in the following pages, of arguably greater importance were developments in other areas of military technology that drove demand for the capability to undertake underwater attacks against increasingly well armed and well protected self-propelled ships capable of enforcing a close blockade.

Whether the *Turtle* in 1776 or Fulton's *Nautilus*, a means of underwater attack was inherently most appealing to the states that did not have the best warships and/or the largest fleet. Above all, underwater attack was appealing to weaker navies facing the threat of blockade and the persons and navies involved with the development of submarines over the nineteenth century therefore tended to also be involved with the development of naval mines. After the Napoleonic Wars, the British navy attained what practically amounted to naval supremacy.[9] At the same time, marine steam propulsion, which was initially restricted to support vessels, was successfully employed on converted sailing ships and, in time, on purpose-built steam-powered warships that offered considerable gains in military capabilities. While primarily utilized as an auxiliary means

[9] Paul M. Kennedy, *The Rise and Fall of British Naval Mastery* (London, UK: Allen Lane, 1976)

of propulsion for several decades – steam-powered ships were primarily reliant on sails until much later in the nineteenth century – marine steam propulsion nevertheless increased the practicality of enforcing a close naval blockade. In the era of sailing ships, blockades were rarely capable of preventing all ships from entering or exiting a port and even squadrons of sailing ships could take advantage of favorable winds to escape the subset of the enemy fleet constituting the first line of blockade. In the case of the French navy, for example, the blockades imposed by the British navy were for the most part blockades in being as the bulk of the British fleet remained in port awaiting news of French warships putting out to sea. With marine steam propulsion, heavily armed warships were better positioned to effectively enforce a close blockade with fewer "leakages," a dynamic first put on display in the Crimean War.

Although marine steam propulsion increased the practicality of enforcing a close blockade, it did not make the blockading ships any less vulnerable to attack. From the 1820s onwards, warships and, to a greater degree given the minimal design constraints in terms of weight and volume, coastal fortifications benefitted from advances in naval armament to which warships were vulnerable. The construction of the French sea-going ironclad warship *Gloire* in 1858, however, initiated a multi-decade competitive dynamic in which militaries pursued ever more heavily armed and armored warships that could withstand considerable damage. To counter the effects of advances in firepower, navies even built more heavily armed and armored ships of their own, and the advances in marine steam propulsion facilitated by this by compensating for the additional weight. In this context, competitive warships became increasingly large and expensive, and militaries searched for ways to counter the latest warships without engaging in an

expensive competition of one-upmanship. This is the context in which a growing number of countries took an interest in naval mines, the context in which self-propelled torpedoes – which were at first self-propelled naval mines – were developed, and the context in which militaries developed a renewed interest in submarines.[10]

In 1863, France, which had quickly fallen behind in the design and construction of ironclad warships, launched the *Plongeur*, the first-self-propelled submarine.[11] Although the *Plongeur* was a very innovative design for its time, it was impractical for military use and therefore had no effects on fielded military capabilities. By the standards of later (self-propelled) submarines developed before the First World War, the *Plongeur* was fairly large and displaced over 350 tonnes. Its size stemmed from the use of a reciprocating compressed air engine as the submarine's sole propulsion system. Compressed air is not a high-density store of energy and much of the submarine's interior was therefore taken up by compressed air tanks. Although self-propelled unlike earlier human-powered submarines, the limitations of the compressed air engine resulted in a slow speed of under 10 km/h and, more importantly, a poor maximum range of just 10 kilometers. For armament, the *Plongeur* was equipped with a single spar torpedo – a naval mine mounted on a pole at the bow of a vessel – that the submarine had to essentially ram into a target ship. As a result, the *Plongeur* could, at best, be employed for the purposes of harbor defense in the context of a close blockade but was otherwise barely seaworthy and impractical against anything other than a stationary ship in much the same manner as the human-powered *Nautilus* that preceded it by some sixty years.

[10] This is also the context in which the *jeune ecole* school of thought developed in France. See Arne Røksund, *The Jeune Ecole: The Strategy of the Weak* (Leiden, Netherlands: Brill, 2007)
[11] "The Model of Le Plongeur," Musee National De La Marine, no date, https://www.musee-marine.fr/en/collections/rochefort/the-model-of-le-plongeur.html

The *Plongeur* remained a one-off proof of concept and did not enter service. Although uncontestably innovative, it is important to recognize that the first self-propelled submarine was developed in a context in which experimental submarines were periodically built in multiple countries. In 1859, for example, a human-powered submarine, the *Ictineo*, was built in Spain. During the American Civil War, the Confederacy developed the *Hunley*, another human-powered submarine, one that recorded the first successful sinking of a surface ship with a submarine (submersible) through the use of a spar torpedo. Notably, the *Hunley* was used against a steam-powered ship that was enforcing a close blockade, and, not unimportantly, the *Hunley* was itself lost in the process. In principle, a self-propelled *Plongeur*-like submarine employing compressed air engines could be employed in a similar manner but in many respects, the *Hunley*'s success was a one-off low probability event and surface ships had ample scope to adapt to counter future attacks. For submarines to be more effective and for the cost of development and manufacturing a submarine to be justified outside the desperate wartime context of an active blockade, designers required a major advance in propulsion technology, one that would not emerge until the 1890s.

While submarine technology remained at an impasse – possible but not practical and therefore ineffectual – militaries and designers turned their attention toward far simpler means of underwater attack. By ca. 1860, the concept of a naval mine was not new. A spar torpedo, the armament of the *Plongeur* and *Hunley*, was, after all, nothing more than a naval mine placed on a pole so that the deploying vessel, which could just as easily be a rowboat or small steam-powered surface vessel, could maintain a safe distance from the explosion. From ca. 1850 onwards, however, militaries worked to create floating

mines to bolster port defenses and deter close blockade. During the Crimean War, Russia bolstered the defenses of coastal fortifications with electrically detonated moored mines, and naval mines were also used in the American Civil War.[12] Mines were, of course, stationary weapons that required an unwittingly cooperative target to function. In the context of this glaring limitation and the growing demand for a means of countering the ever larger, better-armed, and better-armored warships being built, designers set about trying to develop self-propelled naval mines that could move toward a target ship instead of lying in wait. In 1866, such efforts resulted in the development of the Whitehead torpedo, the first self-propelled torpedo.[13]

On paper, the Whitehead self-propelled torpedo was a revolutionary new military technology that constituted a severe threat to surface ships. In principle, the Whitehead torpedo also offered submarines an effective means of underwater attack that could compensate for the limitations of human-powered submarines and submarines powered by compressed air engines. In practice, however, early versions of the Whitehead torpedo were innovative but not very practical and, as such, were very limited in terms of the effects engendered on the fielded military capabilities of adopting states. While its design was progressively improved over several decades, all versions of the Whitehead torpedo in existence until the early twentieth century, which is to say all versions in existence for the better part of forty years, were limited by the design's use of a compressed air engine. Given the inherent limitations of this propulsion system, one also used on the French self-

[12] Sabrina R. Edlow, U.S. Employment of Naval Mines: A Chronology (Alexandria, VA: Center for Naval Analyses, 1997), https://apps.dtic.mil/sti/tr/pdf/ADA362483.pdf

[13] Edwyn Gray, *The Devil's Device: The Story of Robert Whitehead, Inventor of the Torpedo* (Annapolis, MD: Naval Institute Press, 1991); Thomas Wildenberg and Norman Polmar, *Ship Killer: A History of the American Torpedo* (Annapolis, MD: Naval Institute Press, 2010)

propelled submarine *Plongeur*, the earliest Whitehead torpedoes, which were unguided, had a range of no more than two hundred meters, and were also quite slow, being capable of a maximum speed of no more than 10 knots. As a result, the vessel launching the torpedo would be within range of the destructive effects of naval guns and the torpedo could not give chase to a fast-moving let alone a maneuvering target. Early Whitehead torpedoes were also unreliable, a dynamic shaped by the limitations of control systems that would take years of trials to correct. For these reasons, the introduction of the Whitehead torpedo did not engender particularly sizeable gains in the fielded military capabilities of states. With no practical submarines available to employ highly impractical early torpedoes, the introduction of the Whitehead torpedo also had no bearing on submarine technology even as it offered militaries a new approach to underwater attack and constituted a practical application of compressed air engine technology to military purposes. Although the introduction of self-propelled torpedoes by the late 1860s nominally offered submarines a more practical armament than the earlier (non-self-propelled) spar torpedo, it is important to recognize that this technological advance did not drive militaries to pursue submarine technology for at least three interrelated reasons.

First, submarine technology remained limited by a high level of technological immaturity. Innovative though the self-propelled submarine *Plongeur* may have been for its time, a compressed air engine utilized a "fuel" with abysmal energy density and the result was exceedingly limited range and endurance. While a larger submarine could, of course, carry a larger store of compressed air, this was a losing proposition as the weight and volume of the submarine would have to increase in turn, and a crewed submarine,

particularly a submarine carrying multiple unguided torpedoes, was necessarily large and heavy to begin with. In the absence of more substantive advances in propulsion systems for submarines at a time in which there were successive advances in marine steam propulsion for surface ships, submarines with or without torpedoes remained an impractical albeit *possible* military technology. At the same time, compressed air engines were adequate for a functional self-propelled torpedo that was much lighter and smaller than a crewed submarine (i.e., this advance in propulsion technology was adequate for a functional self-propelled torpedo while being woefully inadequate for a self-propelled crewed submarine).

Second, early torpedoes had major limitations irrespective of whether the torpedo was launched from a submarine, a surface warship, or even a shore battery (i.e., coastal defenses). The relatively slow speed and short range of early self-propelled torpedoes, which were wholly unguided, meant that the launch platform had to be exceedingly close – no more than several hundred meters for the first Whitehead torpedoes – from its intended target to have any chance of impacting the target ship (early torpedoes were exclusively equipped with contact-detonating warheads). This was a dangerous undertaking for an armored self-propelled surface warship given the threat posed by warships equipped with rifled guns firing explosive shells and armor-piercing projectiles, but armored self-propelled surface warships could at least realistically maneuver toward a target vessel – while firing at an enemy ship and sustaining fire – so as to bring the target ship within range of a torpedo. In a human-powered submarine or even one powered by a compressed air engine, firing a short-range unguided torpedo at an enemy warship was an

unrealistic proposition other than in highly favorable – and uncommon if not unrealistic[14] – scenarios such as when a submarine would attack a ship lying at anchor or when a submarine was used in a harbor defense role. The increasingly widespread employment of marine steam propulsion, of course, allowed surface ships to maneuver irrespective of wind direction and wind speed and, once detected, early submarines could not hope to maneuver to bring a moving ship into firing position to employ short-range torpedoes.

Third and perhaps most importantly, militaries did not need submarines, whatever their means of locomotion, to employ self-propelled torpedoes, and such torpedoes made ineffective crewed submarines redundant, at least until there was a substantive breakthrough in submarine propulsion. Insofar as the new self-propelled torpedoes offered militaries any appreciable gain in terms of fielded military capabilities, this could be obtained by installing torpedoes on surface warships. While large surface warships, including so-called battleships, were increasingly equipped with torpedoes, by the 1870s, advances in marine steam propulsion allowed for the construction of relatively inexpensive and, when compared to large, armored surface warships, fast motorboats. These vessels – torpedo boats – were far too small to carry guns capable of defeating the armor found on the latest classes of large warships (i.e., cruisers and battleships) but were large enough to carry several torpedoes, each of which held, in principle, the potential of crippling a much larger armored warship.

[14] By unrealistic, I do not mean that such scenarios were implausible; rather, I mean that developing and fielding a new military technology in the hopes that an enemy repeatedly plays into the gambit is unrealistic.

The combination of self-propelled torpedoes with torpedo boats of relatively low cost, small size (i.e., small visual profile),[15] and high speed, limited though the combination was by poor seakeeping, limited range/endurance, and vulnerability to enemy gunfire, allowed militaries to obtain essentially every quality of a torpedo-armed submarine minus the capacity of submergence. Had submarines of this timeframe been capable of sustained and/or high-speed underwater operation, then torpedo boats may have been an inferior substitute for submarines as a means of employing self-propelled torpedoes. Given the immaturity of submarine technology and the fact that submarines of the period were no more than crude submersibles, however, torpedo boats were both practical and effective substitutes for submarines in many of the offensive and defensive roles envisaged for the latter. As such, the advent of the self-propelled torpedo arguably undermined the case for submarines such that the intimate relationship between submarines and torpedoes found in later periods appears anachronistic when viewing the late nineteenth century.

For such reasons, militaries did not adopt submarines to any appreciable degree from the 1860s through the 1880s and the combination of technological immaturity and minimal adoption resulted in periodic advances in submarine technology having negligible implications for the fielded military capabilities of states. By the early 1890s, however, several technological advances intersected and allowed for more practical and potentially more effective submarines, and this led to renewed military interest in submarine technology. Before examining these technological advances – supply-side factors – and the new generation of submarines that these enabled, it is productive to also

[15] I.e., a small target for enemy gunners.

consider demand-side factors that help explain why submarines became more appealing to militaries in the 1890s despite the widespread availability of torpedo boats and other torpedo-equipped surface vessels.

An exclusively supply-side analysis of technological change highlights the opportunities created by technological advances but does not explain why human actors pursued said technological opportunities. As explained earlier in this chapter, throughout the nineteenth century, militaries facing the reality or prospect of blockade by a stronger military looked to means of underwater attack when they could not realistically challenge an enemy fleet's warships on the surface. We can see this dynamic in the case of Fulton's *Nautilus* in France (as well as the submersible *Turtle* from the American Revolutionary War), we can see this dynamic in the history of naval mines, and we can see this dynamic in the 1860s when both the Confederate States of America and France independently pursued new human-powered and self-propelled submarines, respectively. The advent of self-propelled torpedoes, coming as it did alongside the widespread adoption of naval mines, offered militaries a means of more practically accomplishing this long-held objective without crewed submarines of any type. The adoption and diffusion of self-propelled torpedoes, and the construction and deployment of dedicated torpedo boats, however, initiated a measure-countermeasure competition that engendered renewed demand for a vessel capable of submergence and underwater locomotion to attack surface warships.

By the start of the 1890s, torpedoes were increasingly prevalent, being carried on both larger surface warships (including battleships) and much smaller torpedo boats, and many warship types between these extremes. In response to this growing threat to surface

warships, militaries pursued countermeasures to protect their surface ships. Given that self-propelled torpedoes of this period were outranged by the large caliber guns carried by large warships, engaging in a gunnery duel could spare a large warship from a successful torpedo attack carried out by another large warship. Torpedoes were, after all, unguided, and the commanders of large surface ships were likely to prioritize positioning their ship to prevail in a gunnery duel over inflicting a low-probability strike from an unguided torpedo. Small torpedo boats, were, however, more difficult to counter. These small and fast vessels were difficult to detect at sea (the corollary to this was that most torpedo boats had poor seakeeping and were not, as such, a major threat outside of coastal waters) and the slow-firing large-caliber guns required to counter the latest large surface warships at range were unsuited to countering torpedo boats. Stated differently, decades-long trends in warship design emphasizing armor and armament resulted in large warships being equipped with small numbers of progressively larger slow-firing guns capable of defeating the armor found on the latest surface warships. In contrast, to counter torpedo boats, militaries required guns with a much higher rate of fire, and this resulted in the development of small(er) caliber so-called quick-firing[16] naval guns, a new technology that militaries deployed at sea (and in coastal defense roles) some two decades before quick-firing artillery pieces were used by ground formations. At the same time, large surface warships were equipped with torpedo nets for use when a warship was anchored.

The aforementioned countermeasures were not viewed as wholly adequate by then-contemporary militaries, and this was particularly the case in the United Kingdom,

[16] This was the culmination of a series of technological developments in the area of gun design, gun mounts, and (fixed) ammunition.

which, by virtue of having the largest and most powerful surface fleet, had the most to lose if small and comparatively inexpensive torpedo boats could, as many then either feared or hoped, render "obsolete" the large surface warships upon which British naval supremacy rested. To maintain the viability of surface warships in the face of the threat posed by self-propelled torpedoes and torpedo boats, the British military led the way in the development of a new capability set that lay at the intersection of technological and organizational-doctrinal countermeasures, one that shaped the effectiveness and, as such, the implications of later advances in submarine technology.

While the largest warships of a given era tend to receive the most attention, militaries have always employed much far greater numbers of smaller, less expensive vessels for a variety of combat-related roles (i.e., beyond support and logistics roles) whether sloops and brigs (sailing ships) or various types of smaller– and invariably unarmored – vessels equipped with marine steam propulsion. Such vessels were rarely used in battles involving much larger sailing warships (i.e., ships of the line and large frigates) but could be used to screen a fleet and patrol a coastline. With the advent of marine steam propulsion, it became more practical for warships to move together in close formation, and smaller self-propelled warships, what came to be known as torpedo boat destroyers (later simply destroyers, although the term developed a very different connotation from the Cold War onwards), were developed and deployed to defend larger warships from torpedo boats. These torpedo boat destroyers, while considerably larger than the first torpedo boats (by 1914, there was considerable design convergence between what were initially hunter and prey, although this varied by country), were sufficiently fast and maneuverable to counter torpedo boats and were equipped with smaller caliber

armament including quick-firing guns. In time, the torpedo boat destroyers were also equipped with torpedoes of their own to screen against fast light cruisers that operated primarily as scouts for formations of battleships, and could also launch a torpedo attack on other large warships outside of coastal waters (unlike torpedo boats).

By the 1890s, this measure-countermeasure competition between large surface warships (which increasingly carried torpedoes of their own) and smaller torpedo-carrying warships (ranging from large light cruisers to humble coastal torpedo boats) was in full swing and the efficacy of torpedo attack was far from certain notwithstanding the gradual improvements in self-propelled torpedo technology in terms of reliability, range, and speed. Rather paradoxically, torpedo-carrying vessels, large and small, were increasingly vulnerable to gunfire as quick-firing guns outranged torpedoes, and this generated renewed interest in a means of attacking surface warships from beneath the water's surface. Stated differently, measure-countermeasure dynamics between torpedo-carrying surface warships created a situation in which militaries became interested in what amounted to a submersible torpedo boat.

5.1 Late Nineteenth-Century Technological Breakthroughs in Submarine Technology

As noted earlier, all submarines in existence prior to ca. 1860 relied upon the physical exertions of onboard humans as the source of motive power. Over the following three decades or so, a variety of self-propelled submarine designs were proposed, and a few were built that were powered by compressed air (including submarines that were equipped with steam engines that had to be shut down before the submarine could

submerge).[17] Although these were specimens of technological advancement, all such designs were exceedingly limited in terms of practicality and engendered negligible if not non-existent gains in fielded military capabilities for the states that pursued these designs. In 1888, however, France launched a very small – displacing around 30 tonnes – experimental self-propelled submarine, the *Gymnote*, that was driven by an electric motor powered by electric batteries. This was followed by two additional small, experimental submarines in the 1890s – the *Morse* and the *Gustave Zede* – that were also wholly reliant on electric propulsion. Although these were experimental designs – no more than prototypes nominally in military service, existed in militarily insignificant low single-digit numbers, and, by virtue of their small size and poor range, were decidedly limited in terms of operational effectiveness beyond the immediate vicinity of a harbor, the advent of electrically powered self-propelled submarines nevertheless constituted a watershed for submarine technology.

Electric propulsion offered a practical path forward for submarine technology as a means to submerged locomotion, which is the essence of a submarine and what differentiates a submersible vessel from a self-propelled surface vessel including low-profile vessels with superstructures that minimally extend above the waterline (such as some torpedo boat designs). While marine steam propulsion had, over many decades, fundamentally altered the dynamics of naval warfare and maritime commerce, it was wholly unsuited to the purposes of underwater propulsion. Steam engines released far too much heat that had to be contained within a submerged hull of a crewed vessel (which

[17] Rob Dunn, "The Writer Who Built the World's First Engine-Powered Submarine," Smithsonian Magazine, September 17, 2013, https://www.smithsonianmag.com/history/the-writer-who-built-the-worlds-first-engine-powered-submarine-8718089/ ; V. Lansing Collins, "The Sultan's Submarines," *Proceedings* 100, No. 10. (1974).

would remain after the engine was quenched), required a large and continuous supply of oxygen to facilitate combustion (which in a submerged enclosed hull would result in the rapid asphyxiation of the crew), and released harmful combustion byproducts including carbon monoxide which would kill the crew if nothing else did. More generally, steam propulsion remained characterized by low power density (higher power density steam turbines were not developed until the 1890s and would not become widespread until the twentieth century), and coal, the primary fuel utilized for marine steam propulsion until the early twentieth century, was characterized by low energy density when compared with petroleum products, which were not widely utilized for propulsion on land let alone at sea until the 1880s. Militaries needed a new and different means of motive power to develop practical submarines and electric propulsion offered just that.

While the advent of electrically driven submarines powered by batteries heralded the beginning of an era in which more practical submerged vessels became possible, it is important to appreciate that electric battery technology was close to one hundred years old when the first battery-powered submarines were built. Electric batteries were first developed in 1799 (i.e., the voltaic pile), at approximately the same time as Fulton's *Nautilus* and several decades before the discovery of electromagnetic induction. Such early electric batteries were, of course, exceedingly limited in terms of practical applications and, not unimportantly, could not be recharged (i.e., these were primary cells) and this meant that such batteries could not be used to propel a submarine even if a submarine could carry such a (very large) number of batteries as to compensate for the very poor energy density of early batteries. By the 1860s, however, lead-acid batteries were developed, and beyond the improvements in battery chemistry, these new batteries

were secondary cells that could be recharged by electrical generators, but the military implications of this and other advances in battery technology over the nineteenth century were persistently limited by the poor energy density and specific energy of available batteries. With low energy density/low specific energy batteries, a large weight-volume of batteries was required to sustain the movement of a submerged vessel for a meaningful distance in the face of hydrodynamic drag. Given the very small size (enclosed volume and displacement) of submarines in this timeframe, the submerged range of the first electrically-powered submarines was poor – no more than several tens of kilometers at a crawling speed of around 5 km/h (with the energy supply being fixed, range decreased with higher submerged speed).

Given the aforementioned considerations, the advent of electric propulsion was inadequate to enable more practical submarines, and the three experimental (exclusively) electrically powered submarines built by France from the late 1880s onward amounted to little in terms of gains in fielded military capabilities. Notwithstanding the innovation on display, these very small submarines, lacking an onboard electrical generator or means to power one, had to be charged while moored to a pier, could briefly patrol an area no more than several dozen kilometers from said port, and would then have to return to port to have their batteries recharged. In practice, this limited these submarines to the role of harbor defense, one in which such submarines would compete with more practical fast torpedo boats, naval mines, and longer-range coastal artillery as the century drew to a close.

By the 1890s, the practical path toward developing an effective submarine required finding a way to recharge low energy density batteries while a submarine was

surfaced. This way, the submarine could operate as a surface vessel – a surfaced submersible – and only submerge when required to approach or escape an enemy ship or otherwise remain undetected. Marine steam propulsion was, of course, an option for surfaced operation, and steam engines were employed for electricity generation on land, but steam engines could not be immediately shut down to allow the submarine to dive on short notice and more generally posed a hazard to the submarine and its crew. As such, using steam engines to drive electrical generators and charge batteries did not offer a practical way forward, and, absent a more suitable powerplant, militaries were better off pursuing a less complex and less expensive fast torpedo boat that could effectively employ the latest in more compact – higher power density – steam engines.

While submarine technology remained a backwater of technological development into the 1890s, the development of internal combustion engines (steam engines being external combustion engines) from the 1860s onward, particularly Otto-cycle engines using some form of liquid oil (not necessarily a petroleum product) for combustion (Diesel-cycle compression-ignition engines were developed in the 1890s and were limited in terms of practical applications until the 1900s) offered the designers of submarines with new options. Internal combustion engines offered superior power density when compared with even the latest steam engines, including steam turbines (developed in the 1880s), and this was a welcome attribute given the limited space available within a submarine's hull. Moreover, internal combustion engines offered a fast start/stop process and did not require constant human effort to fuel in the manner of stoking a steam engine with coal. This made internal combustion engines an excellent candidate for both propelling a submarine on the surface and recharging its batteries.

By the 1890s, the developmental trajectories of the combination of electric motors and electrical batteries, on the one hand, and internal combustion engines, on the other, converged and allowed for the development of more practical submarine designs. This was, however, a partial, albeit consequential, solution, in that while an internal combustion engine did not release as much waste heat (or steam) as a steam engine and did not need to cool down before a submarine could submerge, running an internal combustion engine while submerged would still rapidly asphyxiate the crew by consuming all available oxygen and poisoning the atmosphere inside the submarine. As a result of this and the enduring limitations of batteries in terms of energy density, submarines that made use of the combination of internal combustion engines and electric motors powered by batteries were necessarily submersibles that primarily operated on the surface – to access atmospheric oxygen and exhaust noxious combustion products – and, when submerged, had to draw upon the limited reserve of electrical energy stored in low energy density/low specific energy electric batteries.

While France led the way in developing and testing (very small) submarines that were exclusively electrically powered and thereafter pursued submarines equipped with steam engines and batteries into the early 1900s, a private venture in the United States, the country at the forefront of the adoption and development of internal combustion engines, led to what became the first practical manifestation of the combination of internal combustion engines, electric motors, and electric batteries for the purposes of submarine propulsion. The individual primarily responsible for the development of submarines equipped with this hybrid combination of propulsion systems, John Holland, had been working on submarines since the mid-1870s and had built and tested several

prototypes, none of which entered military service as a result of major deficiencies. In 1893, Holland was awarded a contract to develop a submersible torpedo boat for the United States navy and this resulted in the *Plunger*, a small steam-powered submarine that was rejected by its intended operator after undergoing testing.[18] Holland's next design, the *Holland VI*, which was launched in 1897, was however more successful and was commissioned by the American navy, and became the first submarine to successfully pair an internal combustion engine with the combination of an electric motor and battery.

Beyond being the first practical self-propelled submarine capable of both surface and subsurface operations beyond the immediate vicinity of a port, Holland's designs were exported to the United Kingdom, which built six *Holland*-class submarines, its first commissioned submarines (from 1903) and Holland's submarines were later also exported to Japan. Innovative though the American (and later British) submarines designed by Holland were, it is important to recognize what these submarines were and were not capable of. Holland's submarines were very small and exceedingly limited in terms of payload, range, and endurance. The first submarine commissioned by the United States, the USS *Holland*, displaced less than 100 tonnes, had a maximum speed of just six knots whilst surfaced, and a maximum range of 200 nautical miles while surfaced and just 30 nautical miles while submerged. On paper, the USS *Holland* carried three torpedoes (with only one torpedo tube) and could, as such, attack surface ships but the torpedoes were unguided and limited in terms of range and, no less importantly, the *Holland*'s crew of six were poorly positioned to find targets to attack and bring the submarine within firing range while surfaced let alone while submerged. Realistically,

[18] Norman Friedman, *U.S. Submarines Through 1945: An Illustrated Design History* (Annapolis, MD: Naval Institute Press, 1995)

such submarines, more accurately characterized as small and slow submersible torpedo boats, were largely limited to harbor defense roles. As a result, this late nineteenth-century technological breakthrough laid the foundations for future, more effectual, technological advances but did not in itself engender sizeable and abrupt shifts in fielded military capabilities.

By 1900, which is to say some one hundred years after Fulton's human-powered *Nautilus* and some forty years after the French self-propelled *Plongeur* were respectively tested, gradual advances in submarine technology allowed for functional, but not very practical or effective, self-propelled submarines. By the end of 1901, three countries, France, the United States, and the United Kingdom, operated submarines (in single-digit quantities) and, as such, expanded the realm of undersea warfare beyond mines and (self-propelled) torpedoes, both of which had been commonplace for several decades. Innovative though the examined efforts toward the development of submarines were, the implications for the fielded military capabilities of states were negligible if not non-existent throughout the nineteenth century.

At the turn of the twentieth century, submarine technology remained highly immature and limited in terms of practicality. The latest self-propelled submarines, including those of the Holland designs, while functional and conducive to the development of experience and learning (including training to defend against attacks by submarines), had a long list of monadic limitations emanating from their small size (which limited the number of torpedoes and the amount of fuel that could be carried), poor seakeeping on the surface, restrictions in terms of underwater navigation, and, not least, the absence of any onboard sensors or communications systems. At first glance,

these small and slow submersible torpedo boats – submarines – presented a new threat to surface warships but the reality was very different not least because the measure-countermeasure competition between torpedo-launching vessels and surface ships was already well underway by the time that practical self-propelled submarines joined the fray. Realistically, such vessels were limited to operations in the environs of a port and adjacent coastal waters and militaries were, all things considered, better served by allocating their finite resources toward torpedo boats and torpedo-carrying destroyers as well as other coastal defenses such as naval mines and coastal artillery. This, to be clear, is not to say that submarine technology had no potential by the turn of the century or that these developmental efforts were a waste of resources. Rather, the point being made is that submarine technology remained in an experimental phase after a hundred or so years of development such that the non-zero shifts engendered in fielded military capabilities of states were decidedly neither sizeable nor abrupt.

A potential objection to the above characterization of the effects of advances in submarine technology on fielded military capabilities in the ca. 1800 to 1900 timeframe is that submarines were not built or used in any appreciable numbers in the nineteenth century. This is, of course, an accurate statement but dwelling on this matter misses the point. The argument made in this chapter is that militaries *could have* built and deployed submarines throughout much of the nineteenth century but did not and had good reasons for not doing so. This was the case in the early 1800s, the 1860s and 1870s, and even in the 1890s. Whatever the shifts engendered by advances in submarine technology in later timeframes, observable technological limitations indicate there is no evidence that the advent and century-long evolution of submarine technology engendered sizeable and

abrupt shifts in the fielded military capabilities of states during the ca. 1800-1900 timeframe. In this respect, the experience with advances in submarine technology over the nineteenth century is typical of the effects of advances in military technology across time, space, and physical domains that tend to be characterized by many years, often decades, of gradual technological development that engender limited effects on the *fielded* military capabilities of states whatever the later potential and realized military effects. A potential counter to this argument could be that the experiences of states in the early twentieth century, that is, in the years leading up to the First World War and during the conflict, suggest that advances in submarine technology did in fact engender significant shifts in the fielded military capabilities of states. As I will explain in the following chapters, however, the interplay of monadic limitations and dyadic vulnerabilities (to countermeasures) meant that the dynamics highlighted in this chapter focusing on the nineteenth century, during which submarine technology was essentially militarily irrelevant, were of enduring relevance over the first half of the twentieth century.

Chapter 6: Submarine Technology ca. 1900-1914

Whereas the implications of advances in armored combat vehicle technology (of the wheeled variety) on land warfare barely elicited debate in the lead-up to the First World War, an extensive debate transpired about the implications of advances in submarine technology, particularly the pairing of new submarine designs with new types of torpedoes, in the period ca. 1900 to 1914. Although the United Kingdom, France, and the United States built and fielded sizeable "fleets" of what were, all things considered, very small coastal submarines, at the start of the century, there were as late as 1910 relatively few of the progressively larger and more capable submarine designs in existence let alone in active military service. As a result, such debates about the implications of submarine technology were mostly theoretical and focused on prospective technological developments rather than then contemporaneous fielded military capabilities. This dynamic changed by 1914 as a result of a surge in submarine construction in the five or so years leading up to the First World War, but the implication is that there was considerable anticipation of the implications of advances in submarine technology before such advances were realized in the form of floating submarines and before such submarines were fielded in significant numbers.

The next chapter will examine the implications of advances in submarine technology during the First World War and thereby determine whether there was a gap between expectation and reality about the military effects of advances in submarine technology in the early twentieth century. To anchor the next chapter's examination of wartime performance, this chapter focuses on the implications of the advances in submarine technology made in the fourteen or so year-long period between the advent of

more practical submarines ca. 1900, that were equipped with both internal combustion engines and the combination of electric motors and electric batteries, and the start of the First World War. Did the advances in submarine technology made in this timeframe engender sizeable and abrupt shifts in the fielded military capabilities of the states?

Many writings on submarine technology, including scholarship that examines submarine technology as a case of technological change, do not longitudinally examine advances in submarine technology. Instead, many writings offer a cursory examination of technological developments in the lead-up to 1914 and focus on wartime observables from the First World War and the Second World War (separated by a typically cursory examination of technological developments that transpired during the Interwar Period). This is, of course, understandable given that although submarines were deployed by militaries in increasing numbers from ca. 1900 onwards, we have relatively little in the way of empirics through which to discern the military implications of the fielding of submarines. Lacking empirics of combat employment through which to discern the implications of advances in submarine technology on the fielded military capabilities of states, we must turn to the analytical framework developed in Chapter Three to develop a better appreciation of what submarine technology did and did not offer states in terms of fielded military capabilities in the years leading up to the First World War.

6.1 Examining the Military Effects of Technological Change Through a Technology Development Lens

Irrespective of how much – or how little – advances in submarine technology offered in terms of capability gains in the ca. 1900-1914 timeframe, it is important to focus on the

effects of technological change on the *fielded* military capabilities of states and this requires us to view technology development as a process that begins with research and development efforts and thereafter continues to testing, manufacturing, and, finally, the fielding of a new technology in militarily significant numbers.

Although submarines were much smaller than the largest surface warship types that leading militaries built in large numbers in the fourteen or so years leading up to the First World War, submarines were nevertheless expensive to manufacture and were, for their time, complex forms of machinery to build and operate. At the start of the twentieth century, there was very little experience with the design and construction of submarines among designers let alone production facilities and the broader industrial supply chain, and it took much effort and time to accumulate experience and refine submarine designs so as to reduce the cost and complexity of manufacture. In contrast, at the turn of the century, militaries had decades of experience with the design and construction of steam-powered ships, armored warships, and iron-hulled ships and the result was that militaries were poorly positioned to rapidly deploy large numbers of submarines even if they were so inclined. Compounding matters, the first submarine designs available in the early 1900s were, as I will explain in the following pages, exceedingly limited in terms of performance and military utility, and successively improved designs could be developed faster than the prior design(s) could be built and fielded. As a result, early adopters of the latest in submarine technology in the late 1890s and early 1900s, namely France, the United States, and the United Kingdom, found themselves operating large numbers of older, smaller, and less capable coastal submarines that were of very limited and fast diminishing military utility by ca. 1910 let alone 1914. Germany, a relatively late adopter

(i.e., a state that emulated advances in submarine technology made by others), did not build large numbers of very small submarines in the same manner as France, the United States, and the United Kingdom, but its (successful) efforts to quickly reach the forefront of submarine technology on the basis of its internationally competitive industries resulted in the building of many different submarine designs (i.e., many classes, or distinct designs, each composed of a handful of vessels) that, as a collective, made for an excellent naval technology exhibition but a poor fleet to operate and sustain. Leading militaries therefore found themselves in a broadly similar position by the start of the First World War – the capabilities offered by their decidedly heterogenous – hybrid – submarine fleets were less than the sum of the parts as there was tremendous variance between the military utility of their latest and oldest submarines even though, as of ca. 1910, the vast majority of submarines in existence was no more than a decade old.

While submarine technology advanced considerably in the fourteen years leading up to the First World War on the heels of a surge in investment worldwide directed toward naval capabilities, militaries did not have infinite resources to allocate toward naval construction. Having been built at considerable expense, essentially every submarine built since the late 1890s that was not lost in an accident remained in service going into 1914. Although increasingly relegated to secondary roles such as harbor defense and training support, militaries did not have the resources to maintain force structure (i.e., fleet size), however nominal, while simultaneously fielding only the latest submarine designs which tended to be considerably more expensive and complex to build and operate given the increased size and displacement of these designs. The capability gains offered by advances in submarine technology made over the ca. 1900-1914

timeframe were not, in other words, free, and as early as the start of the first decade of the twentieth century, militaries had to engage in a delicate balancing act of maintaining increasingly heterogeneous hybrid force structures until enough new submarines reflecting the latest technological advances could be built to avoid cutting back on force structure as a cost of modernization.

In a world of infinite resources and no trade-offs in resource allocation, militaries could have replaced their existing submarines with the latest designs at a rapid pace. Instead, the effects of technological change were moderated by the fact that the development of a new military technology has zero effects on the fielded military capabilities of states until the technology is tested, manufactured, and fielded in militarily significant numbers. In the case of submarine technology in the period ca. 1900 to 1914, the fielded military capabilities of states lagged the nominal capability gains engendered by technological change. For this reason, it is important not to conflate the speed at which submarine technology advanced in this timeframe with shifts engendered in the (real-world) *fielded* military capabilities of states as the latter necessarily lagged behind the former.

6.2 Advances in Submarine Technology

Although the military effects of technological change on fielded military capabilities lagged behind the expanding technological frontier, this does not characterize the monadic capability gains that advances in submarine technology offered states. On paper, submarine technology experienced considerable technological development from the years ca. 1900 through 1914 on the basis of accumulating knowledge and experience in

designing, manufacturing, and operating submarines. From 1910 onwards, militaries were building ever greater numbers of progressively larger and more capable submarines, often displacing 400 or so tonnes if not more, that were better suited for operations in open waters (i.e., not restricted to coastal waters). Initially equipped for the most part with internal combustion engines using gasoline or kerosene (alongside an electric motor and electric battery), by 1910 newly built submarines increasingly made use of diesel engines (alongside an electric motor and electric battery). The engines employed on submarines became increasingly reliable and diesel (cycle) engines in particular, also known as compression-ignition engines, used a less volatile fuel than gasoline that allowed for safer operation. With Diesel-cycle engines being far more efficient than Otto-cycle gasoline engines, the new, larger submarine designs equipped with diesel engines benefitted from disproportionately greater range, when transiting on the surface, than preceding designs. The larger hulls of new submarine designs not only allowed for greater fuel storage but also space for additional supplies and crew accommodations, thereby increasing endurance (while surfaced) and facilitating sustained operations at greater distances. Larger submarine hulls also allowed for multiple torpedo tubes and the stowage of reloads, thereby correcting the glaring deficiencies of the very small submarines built and deployed in the early 1900s.

In principle, successive advances in submarine technology made submarines increasingly practical and, as such, offered states substantial gains in their fielded capabilities. In practice, however, the implications of advances in submarine technology in the ca. 1900-1914 timeframe were moderated by enduring technological limitations above all in the area of propulsion technology that meant that even the latest submarines

were no more than submersibles that primarily operated on the surface. As a result, while submarine technology advanced from mere prototypes to practical ocean-going designs in little more than fifteen years, the net effects on the balance of fielded military capabilities were considerably more limited.

For example, while the application of diesel engines to submarines offered states major gains in capabilities, this only improved the qualities of a submarine as a submersible operating on the surface. Battery technology, in contrast, experienced very limited advancement in the ca. 1900-1914 timeframe, and larger submarines therefore only carried more battery cells that remained comparably poor in terms of specific energy and energy density. As a result, by ca. 1913 the latest large submarines had a maximum range of 15,000 kilometers when operating on the surface (while operating at a high cruise speed) but only 150 kilometers when submerged (while operating at a very slow speed). Given the growing chasm between surfaced and submerged range and endurance, even the latest pre-1914 submarines were submersibles that primarily operated on the surface. Operating on the surface, however, meant that a submarine was likely to be detected and likely to be subject to attack by gunfire. A submarine could, of course, dive to evade gunfire but with such poor underwater endurance, submarine crews were incentivized to remain submerged and move at a very slow speed during daylight in high-risk areas and only operate on the surface at speed – and recharge batteries – at night and/or in low visibility conditions or, more generally, in lower threat environments.

While enduring limitations in propulsion technology restricted submarines to primarily operating on the surface, the capabilities that submarines offered states improved as a result of advances in torpedo technology. Until the early 1900s, self-

propelled torpedoes employed compressed air engines, and this limited the speed-range of torpedoes. Given that torpedoes were unguided, greater speed could contribute to improving the accuracy of torpedoes against moving targets by reducing the demands on fire control. An increase in range without an increase in speed, however, would only increase the time to target and thereby reduce the accuracy of unguided torpedoes against moving targets. Over the ca. 1900-1914 timeframe, militaries transitioned from compressed air torpedoes to newer designs that made use of wet heater combustion engines that offered major improvements in terms of speed-range (an approximately 50% increase in range at low-speed settings and an approximately 50% increase in speed at the maximum range of preceding torpedoes).

While advances in torpedo technology offered states gains in fielded capabilities, it is important to recognize that these capability gains were not restricted to submarines and also benefitted torpedo boats and torpedo-armed destroyers. As a result, submarine technology received renewed competition for funding and roles from other types of naval vessels that could substitute for submarines in the employment of wet-heater torpedoes. The latest torpedo boats could not, of course, operate far out to sea but, by virtue of high surface speeds, were far better suited to many coastal defense roles than submarines that could never exist in large enough numbers to compensate for a slow submerged speed that did not allow for short notice deployments in response to the activities of adversary surface ships. Similarly, the latest ocean-going destroyers and light cruisers could employ wet-heater torpedoes while screening a formation of battleships and heavy cruisers, thereby allowing militaries to augment the latest in naval gun technology with the latest in torpedo technology. Not having a monopoly on the employment of wet-heater

torpedoes, the latest submarines were unable to supplant even the lowest tiers of surface warships, and militaries were disincentivized from heavily investing in the latest submarines, thereby moderating the effects of technological change on the fielded military capabilities of states.

6.3 The development of countermeasures for use against submarines

Although there was considerable debate about the implications of advances in submarine technology in the period ca. 1900 to 1914, there is limited evidence of *direct* efforts toward technological and/or organizational-doctrinal countermeasures against submarines in this timeframe. Taken at face value, this may suggest that militaries either did not attempt to counter submarines or that their efforts did not bear fruit, but such interpretations overlook an important consideration – that the threat posed to surface warships by the submersibles of the era was hardly distinguishable from the threat posed by (self-propelled) torpedoes independent of the type of launching vessel and the threat posed by (coastal) torpedo boats and (sea-going) torpedo-armed destroyers in particular. As a result, efforts to counter submarines had considerable overlap with efforts to counter torpedoes launched from surface vessels of all shapes and sizes.

As explained in the previous chapter, militaries adopted two primary countermeasures to the threat posed by torpedo boats – (A) warships were equipped with quick-firing guns to engage an enemy vessel before it launched its unguided torpedoes and (B) militaries protected their large surface warships with a screen of smaller and faster escorts such as destroyers (torpedo boat destroyers). While a submarine that caught a (moving) surface ship off guard in a close-range ambush-type situation had a fair

chance of severely damaging if not sinking the ship, such opportunities were rare, and submarines usually had to position themselves to intercept a moving vessel and bring their torpedoes within range. As a result of the limitations of battery technology – the slow speed and poor range of submerged submarines, submarines were forced to operate on the surface to intercept a moving ship. As a result, surface warships, which were the primary if not the only target of submarines (as determined by adopting states and their many idiosyncrasies) in the pre-1914 period, could practically employ the countermeasures developed for use against torpedo boats against submarines closing in at a relatively high speed whilst surfaced. As for the threat posed by submerged submarines, warships were unlikely to be hit by unguided torpedoes other than in chance ambush-type encounters for as long as the warship was moving and maintained lookouts to detect the tell-tale wake of an active torpedo and thereafter undertook evasive maneuvers in response. Surface warships would, of course, invariably cast anchor, and the largest warships – battleships – received anti-torpedo nets to protect against torpedo attack from a submerged submarine while torpedo-gun boats and destroyers, among other vessels, could patrol the approaches to ports and thereby deter submarines from lying in wait on the surface to dive and intercept a detected warship while submerged.

All things considered, there was a substantial gap between perceptions and expectations of the effects of submarine technology in the years leading up to 1914 and the reality of what progressively improved then-extant submarine technology could and could not offer states in terms of fielded military capabilities. As explained earlier, while submarine technology advanced fairly quickly on the heels of accumulating experience, fielded capabilities lagged as it not only took time for technological advances to be

realized but also tested, built, and deployed in militarily significant numbers. Paradoxically, the fast pace of technological growth, which is reflective of base effects and the very low baseline that self-propelled submarine technology started at by the turn of the century, made the situation more difficult for militaries. In a world of finite resources, states invariably operate heterogenous and hybrid force structures composed of both older and newer equipment. As late as 1914, this meant that whatever the capabilities offered of the latest submarine designs, no state operated a fleet of submarines in which anything close to a majority was composed of the latest designs. Fielded capabilities, therefore, lagged considerably behind the technological frontier. On paper, submarine technology experienced considerable monadic gains in performance growth in the ca. 1900-1914 timeframe. Larger, more capable, and more refined submarines could operate over longer distances and carry more torpedoes. There were, however, also monadic limitations including enduring technological limitations that constrained the real-world implications of submarine technology on the fielded military capabilities of states. In sum, advances in submarine technology in the ca. 1900-1914 timeframe did not engender sizeable and abrupt shifts in the fielded military capabilities of states. As late as 1914, there were not many militarily significant submarines, those that existed had major monadic limitations when employed against both surface warships and merchant shipping, and submarines were also intrinsically vulnerable to early efforts toward anti-submarine countermeasures.

Chapter 7: Submarine Technology ca. 1914-1918

Prior to 1914, submarine technology was untested in combat but the First World War both raised submarines to unexpected heights of significance and subjected a fledgling area of military technology to an unexpectedly intense measure-countermeasure competition. To develop an accurate and holistic appreciation of the military implications of the employment of submarines in the First World War and thereby determine whether experiences with submarine technology support the notion that technological change is a suitable candidate for a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states, we must first recognize that there was not a single experience with submarines in the First World War. There was, as such, not a single set of implications, broadly defined, of the employment of submarines, but many, and these varied with both the military-geographical context of submarine operations and specific dyad-level operational dynamics that were in turn shaped by the variant roles for which submarines were employed. As a result, this chapter will offer a decidedly synthetic analysis of the implications of submarine technology in the First World War timeframe rather than an analysis focusing on the military and diplomatic effects of Germany's employment of submarines to attack merchant shipping or the other, less prominent and less well known, experiences with submarine operations in the First World War. If technological change is a suitable candidate for the role of a prime mover engendering shifts in the fielded military capabilities of states that are both sizeable and abrupt, then we should find evidence of this in the important area of submarine technology in the context of the First World War, the first conflict in which this "new" military technology was employed in combat.

7.1 Wartime advances in submarine technology

In principle, wars create incentives for large-scale research and development efforts in pursuit of more capable armaments that can result in technological breakthroughs. Once developed to a state ready for manufacturing, this can result in substantially improved armaments that engender significant gains in the fielded military capabilities of states. In practice, such an outcome is not guaranteed and is most likely to transpire in (A) wars that are protracted – the research and development, testing, and manufacturing processes tend to take many months if not several years before militarily significant numbers of service-ready specimens of a substantially new technology are available for the fielding process to begin – and (B) in areas of military technology characterized by relatively short lead times for developing and producing a substantially improved clean sheet design.

Notwithstanding the limitations of submarine technology, by ca. 1910 submarines were among the most complex "individual" articles of human technology in existence and the latest large ocean-going submarines regularly took somewhere in the range of twelve to twenty-four months to transition from the laying of the keel in a shipyard to a vessel ready for a new crew. While existing submarine designs could, in principle, be produced more quickly than new clean sheet designs with which production facilities were by definition unfamiliar, the unforgiving underwater operating environment meant that militaries could not simply tinker with an existing design by, for example, swapping out the engine, elongating the hull, or adding many tonnes of equipment without undertaking considerable developmental effort – assuming considerable engineering risk – and subjecting the new submarine to extensive testing. Simply put, a greatly altered

existing submarine design was, in effect, a new submarine design, and the lead times for developing a new submarine were greater than what was required to develop new small arms, new artillery pieces, or even new fixed-wing aircraft. As such, while wars tend to create incentives for the establishment of well-resourced large-scale research and development efforts in pursuit of more capable armaments, submarine technology was an unlikely candidate for a technological breakthrough even in a conflict that, with hindsight, continued for four years and was characterized by unprecedented levels of resource allocation towards military production.

Although the First World War was uncontestably protracted in retrospect, it is important to recognize that states did not know this in advance and therefore pursued the conflict on a day-to-day basis. From the start of the war, states had to balance between effectively husbanding their resources in preparation for a potentially protracted conflict of unknown duration while allocating available resources to replace losses and bolster fielded forces so as to stave off immediate defeat (and ideally snatch an early victory). As a result, from August 1914 onward, states worked to produce as much military equipment as possible and devoted considerable effort toward simplifying existing designs to maximize production output and thereby both compensate for losses and arm freshly raised forces. In this context, ambitious research and developmental efforts that were unlikely to offer capability gains in the near term were generally placed on the backburner until severe rates of attrition and a seemingly never-ending stalemate rekindled interest in new, more capable, designs. As a result, while the production output of submarines increased early in the war, new designs offering substantial improvements in capabilities, apart from those already under development before the start of the war,

were generally unavailable until 1916, were not available in sizeable numbers until 1917 if not 1918 and some would not enter service in militarily significant numbers by November 1918.

When compared with submarines from the early 1900s, ocean-going submarines built in the First World War were very different, being far larger and more capable. When compared with submarines built from circa 1910 onward, however, the submarines built in the First World War were largely more of the same. New designs were of course refined on the basis of accumulating experience, including hitherto non-existent combat experience, and new ocean-going submarines built during the war were generally larger, had greater range and endurance, and carried a greater number of torpedoes. Importantly, First World War submarines were increasingly equipped with radios. Radio(telegraphy) technology, a late nineteenth-century invention that was nearly twenty years old in 1914, nominally allowed (surfaced) submarines to communicate with one another, with ships, and with land-based headquarters and, as such, made submarines more effective and thereby engendered considerable gains in the fielded capabilities of states that operated radio-equipped submarines. Nevertheless, it is important to recognize that militaries were experimenting with radios on submarines before the turn of the decade on the heels growing ubiquity of radios on surface ships. Radios were not, in other words, introduced to submarines during the First World War and did not constitute a technological breakthrough even as the war expedited the employment of radio in submarine fleets and across militaries more generally.

Beyond radios, which became standard equipment on new-build ocean-going submarines over the war, there was little in the way of new technology on most

submarines built during the First World War and this is unsurprising given that there were no fundamental breakthroughs in the component technologies that make up a submarine. Leaving aside the development of progressively larger and better armed, but otherwise not radically different, submarines better suited to long-range, long-endurance commerce raiding operations that were primarily developed and deployed by Germany, two areas of technological development, both also primarily pursued by Germany, are worth considering. By more closely examining the implications of the adoption of new armaments for submarines as well as the development of a new generation of coastal submarines, including coastal minelaying submarines, we can develop a better appreciation of the limited capability gains offered by wartime advances in submarine technology and better understand the degree to which the engendered effects on fielded capabilities were contingent upon the interplay of a very specific military-geographic context and a specific set of priority military objectives in that operational context.

7.2 New Armaments for Submarines: A Primarily German Experience

Although there were no fundamental breakthroughs in submarine technology during the First World War, most larger ocean-going submarines in existence, particularly those operated by Germany, received a new, or at least different, type of armament by the end of the war. Whereas submarines were primarily conceived of as a means of attacking surface warships prior to the First World War, the requirements of commerce raiding, which were, for the most part, exclusive to Germany, drove the adoption of deck-mounted naval guns on submarines and, as such, resulted in submarines being unexpectedly employed as (submersible) gunboats whilst operating on the surface.

Unguided torpedoes were, as previously explained, inaccurate against moving targets, and most torpedoes employed in wartime unsurprising missed their intended targets. With First World War submarines typically carrying fewer than a dozen torpedoes (the smallest submarines typically carried no more than four torpedoes), submarines either required more accurate torpedoes, which did not exist, or a different form of armament in order to attack surface ships more effectively. With submarines primarily operating on the surface as a result of the enduring limitations of battery technology, a surfaced submarine was inherently vulnerable to gunfire from enemy vessels before it could fire its unguided torpedoes with any degree of accuracy. First World War submarines were, as such, not very effective against surface warships unless a submarine was submerged and a submerged submarine was too slow and too limited in terms of submerged speed and range to realistically intercept a moving ship and was therefore largely limited to chance, ambush type encounters.

In Germany, the glaring limitations of what practically amounted to submersible torpedo boats primarily operating on the surface intersected with the inability of the German surface fleet, in many respects the second-most powerful fleet in the world, to readily operate within the North Sea let alone access the world ocean. In this context, not only was the German surface fleet forced to operate as a fleet in being against the British navy, but German cruisers based in Germany could not operate as surface commerce raiders on the high seas and thereby realize the decades-long latent threat to merchant ships on the high seas posed by fast steam-powered cruisers. Moreover, given the geography of the North Sea and the degree to which it was surveilled by British radio-equipped surface vessels, submarines, aircraft (both fixed-wing and lighter-than-air) as

well land-based signals intelligence capabilities, German surface warships, including fast light cruisers, destroyers, and torpedo boats, could not effectively undertake high-speed short duration raids along the British coastline without encountering concerted resistance.

In this military-geographical context and, not least, in the context of a war that was protracted largely as a result of stalemate on the Western Front, a dynamic that was wholly unrelated to any aspect of naval technology, Germany turned to its submarine force to undertake a variety of roles that its surface fleet could not and for which Germany, like every other belligerent, had given exceedingly little attention to prior to the war. Faced with an adversary alliance with greater industrial capacity and unparalleled access to natural resources that Germany could not hope to match, the German navy increasingly turned to commerce raiding, given the country's unique military-geographical context and the effective blockade imposed by the British surface fleet, submarines became Germany's primary instrument of commerce raiding by default. While submarines inherently constituted a means of engaging in commerce raiding in the face of enemy blockade, most of the submarines available in the 1914-1915 timeframe were, by virtue of limited endurance and a limited torpedo load, however, poorly suited to commerce raiding and were therefore poor if necessary substitutes for fast ocean-going cruisers.

While larger submarines could carry additional torpedo reloads and be equipped with a larger number of torpedo tubes, this would require the production of the latest designs as well as new, clean-sheet designs that had to be developed, tested, built, and fielded. Submarines better suited to long-distance commerce raiding missions were not available to Germany in militarily significant numbers until 1916 and only became

common on operational patrols from 1917 onward. With few torpedoes (numbering in the single digits) carried by most submarines operational in the 1914-1915 timeframe, with most torpedoes missing their targets, and with the employment of torpedoes giving away the presence of the submarine and thereby forcing it to dive and expend scarce battery charge (if it was not already submerged) if an enemy warship was in the area, submarine crews had to be exceptionally judicious in the employment of torpedoes. In practice, this meant that non-reckless submarine crews had to forgo attacking detected targets if their torpedoes were unlikely to hit the target or if the target, whether a surface ship, merchant ship, or groups thereof, was well positioned to counterattack. As a result, German submarines were subject to the effects of virtual attrition even in the absence of the sustained development and employment of countermeasures by their adversaries. If submarine crews were less risk averse and instead more reckless, their submarines were more likely to be damaged if not destroyed, and, no less importantly, any submarine that survived high-risk and/or low-probability engagements involving the expenditure of scarce torpedoes would have to cut short its patrol to return to port to rearm. With diesel-powered submarines primarily limited in terms of endurance by supplies and the number of torpedoes carried, submarines needed a new type of armament, and militaries, above all the German military, undertook what was in many respects a regressive technological development – the mounting of deck guns on larger submarines.

As submarines became larger over the 1900s, submarines increasingly operated on the surface and militaries therefore began to experiment with equipping submarines with deck guns of various sizes. The larger the submarine, the more readily it could be equipped with one or two deck guns in the 70-100 mm bore diameter medium caliber

guns that were, in more elaborate turreted form, secondary (if not tertiary) armament on the largest warships and the primary gun-type armament on destroyers. A large submarine displacing over 500 tonnes could readily carry one hundred or more rounds of ammunition for such a deck gun even as it was limited to less than a dozen torpedoes. As German commerce raiding efforts took off in early 1915, German submarines were increasingly equipped with deck guns as an expedient measure, and, in time, newly built submarines received deck guns as standard equipment.

On paper, a submarine equipped with a deck gun had firepower comparable to a small destroyer. Submarines were, however, poor gunnery platforms and one or two medium caliber guns in the 70-100 mm bore diameter range were inadequate for the targeting of surface warships given the absence of the range-finding equipment found on larger surface warships (i.e., deck guns had to be fired at close range in as low trajectory direct-fire weapons). Being unarmored and reliant on the integrity of its hull to withstand hydrostatic pressure while submerged, a surfaced submarine was also in no position to participate in a gunnery duel. One or two deck guns nevertheless allowed submarines to effectively attack most merchant ships as well as coastal vessels (fishing boats, coasters, etc.) given that most First World War era merchant ships were far smaller and slower than late twentieth century let alone twenty-first-century merchant ships. While medium caliber guns were generally inadequate to sink merchant ships, the merchant ship's crew were more likely to surrender after being subject to gunfire and this allowed a submarine to despatch a boarding party (on a rowboat) to scuttle the merchant ship after the crew were allowed to board lifeboats in accordance with the pre-war norms of cruise rules.

In principle, deck guns improved the capabilities offered by submarines in commerce raiding roles, and deck guns played a prominent if largely uncelebrated role in the early successes of the German commerce raiding campaign (alongside the use of boarding parties to scuttle surrendered merchant ships). It is nevertheless important to recognize that deck guns could only be effective for as long as merchant ships were unarmed and unescorted. Before the start of the First World War, leading militaries viewed their country's fastest civilian ships – ocean liners – as auxiliary cruisers that could be used to counter (surface) auxiliary commerce raiders, and some ocean liners were designed from the outset to mount naval guns in times of war. There was, as such, precedent for the arming of merchant ships in the age of marine steam propulsion let alone the age of sail. As the threat posed by highly vulnerable surfaced submarines operating deck guns increased, Germany's enemies were incentivized to arm merchant ships that, when equipped with modest medium-caliber deck guns comparable to those mounted on a submarine, could fire at a detected submarine, and thereby force the submarine to dive and employ one or more of its scarce unguided torpedoes or otherwise risk major damage to the pressure hull. As a result, the equipping of submarines with deck guns was an expedient measure that was only practical against a specific type of target – an unarmed vessel. When faced with all other types of potential targets, including merchant ships escorted by better-armed surface warships, submarines were forced to rely upon the finite number of inaccurate unguided torpedoes available onboard and, as the small supply of torpedoes was exhausted, submarines were forced to cut short their patrols and return to port. The adoption of a new form of armament, deck guns, did not, therefore engender particularly sizeable gains in fielded military capabilities and should

be viewed as a wartime expedient to compensate for the enduring limitations of submarine technology.

7.3 New Coastal Submarines: A Primarily German Experience

Although most new larger and more heavily armed ocean-going submarine designs were not ready for service until the last two years of the First World War, Germany successfully developed and built very large numbers of a series of much smaller, less complex, and less expensive coastal submarines that initially displaced no more than 150 tonnes and, in their final late-war forms, displaced under 600 tonnes (at a time when the latest German ocean-going submarines were exceeding 1,000 tonnes). Submarines of such size were limited in terms of endurance and were not therefore well-suited for long-range operations around the British Isles from German ports along the North Sea. As a result of this dynamic (and the country's early lead in the adoption of diesel propulsion for submarines), Germany had the largest force of longer-range non-coastal submarines at the start of the war.

With Germany securing control over much of the Belgian coastline including the ports of Bruges and Antwerp by the fall of 1914, the German military could pursue the development of new, much smaller submarine designs such as the Type UB I class, a very small coastal submarine that was built in sections at shipyards in Germany, transferred by rail, and reassembled in occupied Belgian ports. The UB I class submarines, of which twenty were built and which began to enter service in 1915,

displaced less than 150 tonnes and could only carry two torpedoes.[19] These constituted, as such, a regressive technological development by the standards of the German shipbuilding industry but nevertheless offered the German military the ability to patrol the shallow coastal waters to the southeast of the British Isles with smaller, less complex, and less expensive submarines that could be quickly built in far greater numbers. As such, wartime changes in the military-geographical context in which submarines operated affected the practicality and efficacy of what were by any standard very modest submarines. The German Type UB I class was followed by the much larger Type UB II and, in turn, the even larger Type UB III, which entered service in 1917 and displaced around 550 tonnes.[20] Submarines of such size could not be built in sections and sent to Belgium by rail for assembly and Germany therefore built these submarines in its own ports, after which these submarines had to transit the North Sea to reach Belgian ports.

This access to Belgian ports, which made much of the British coastline far more proximate to German naval bases, and the new focus on commerce raiding allowed Germany to undertake conceptual innovations in submarine technology. First World War era submarines were, as explained, limited to carrying a very small number of inaccurate unguided torpedoes. A torpedo was, however, no more destructive than an unguided naval mine and First World War surface ships had no means of detecting submerged naval mines other than by visual detection. While extensive defensive minefields were laid by small surface vessels during the First World War, surface vessels were impractical for use in mining an enemy's coastline or the approaches to enemy ports. Submarines were, however, excellent candidates for the role of a covert minelayer. Germany

[19] Gudmunur Helgason, "Type UB I," Uboat.net, no date, https://uboat.net/wwi/types/?type=UB+I
[20] Gudmunur Helgason, "Type UB III," Uboat.net, no date, https://uboat.net/wwi/types/?type=UB+III

therefore developed the Type UC I submarine, which displaced under 200 tonnes and first entered service in 1915.[21] While capable of deploying twelve mines, the Type UC I carried no torpedoes and was followed by the much larger Type UC II class, sixty-four of which were built and entered service from 1916 onward.[22] Displacing under 450 tonnes, this was a much larger submarine and could carry eighteen mines alongside seven torpedoes (in three tubes) and was followed by the broadly similar Type UC III class.[23]

Although minelaying submarines were for the most part a wartime conceptual innovation, it is important to recognize that Germany's coastal submarines and coastal minelaying submarines did not reflect any breakthrough in submarine technology. The only reason that the aforementioned submarine designs became appealing to the German military was because the war – specifically the war on land – changed the military-geographical context in which the German fleet operated. Prior to the First World War, Germany led the way in pursuing larger, longer-range submarines with the endurance required to operate across the North Sea and around the British Isles as a result of its peacetime military-geographical context. With access to ports on the Belgian coastline, however, the German military and shipbuilding industry adapted and concurrently designed and built new coastal submarines alongside new long-range ocean-going submarines to serve as commerce raiders.

It is striking that no other country, including the United Kingdom, followed the German lead across various areas of submarine technology during the First World War. The reasons for this are, of course, straightforward. No country shared Germany's

[21] Gudmunur Helgason, "Type UC I," Uboat.net, no date, https://uboat.net/wwi/types/?type=UC+I
[22] Gudmunur Helgason, "Type UC II," Uboat.net, no date, https://uboat.net/wwi/types/?type=UC+II
[23] Gudmunur Helgason, "Type UC III," Uboat.net, no date, https://uboat.net/wwi/types/?type=UC+III

military-geographical context, which resulted in the world's second-largest surface fleet being bottled in the North Sea, no country was as reliant on its submarine force, and no country shared Germany's pressing need for commerce raiders that could circumvent the British chokehold on the German fleet in the hopes of interdicting supplies and thereby weakening the United Kingdom and France on the Western Front. With the German merchant fleet dispersed in neutral ports, in German ports, or captured by the United Kingdom, and more generally driven from the high seas, the British navy, the world's largest and most capable, had no need for submarines optimized for coastal minelaying or long-range commerce raiding – the German coastline was not a target-rich environment. Instead, British submarines were employed throughout the North Sea to serve as a deterrent against the German surface fleet, deployed to the Baltic Sea on small-scaler commerce raiding operations directed against merchant ships primarily transferring Swedish iron ore to Germany, and, uniquely, employed submarines to attack German (surfaced) submarines entering and exiting ports (i.e., British submarines were uniquely employed as a form of anti-submarine capability). These roles did not require much in the way of conceptual or technological innovation, and, in the First World War, it was largely Germany alone that pursued progressively better adapted, and in that respect more capable, submarines even in the absence of fundamental technological breakthroughs.

All things considered, wartime developments in submarine technology did not constitute a fundamental breakthrough, and the primary performance gains enjoyed by states were manifested in terms of surfaced range and endurance, as well as torpedo load, and gains in these areas strongly correlated with the size, complexity, and cost of submarines that were otherwise largely indistinguishable from pre-1914 submarine

designs. The real-world effects of these areas of performance growth were, not unimportantly, strongly shaped by the unique military-geographical context that Germany found itself in during the war as well as its decision to employ submarines as the primary tool of a far-reaching campaign against maritime commerce. With access to ports in occupied Belgium, Germany was able to undertake a submarine-centric commerce raiding campaign on a scale that would likely have been impractical if it was forced to solely rely upon its much smaller fleet of larger, more expensive longer-range ocean-going submarines operating from German ports. Such larger submarines, which became available in increasing numbers from 1916 onwards, were not, however, intrinsically more capable than smaller coastal submarines. If technological change offered states considerable gains in combat capabilities, then we should find evidence of similarly sized and similarly expensive submarines becoming more capable and/or a widening gap between the performance of the latest large ocean-going submarines with the latest smaller coastal submarines. There is, however, no evidence of such a dynamic and there is, as such, little support for the proposition that wartime advances in submarine technology engendered sizeable and abrupt shifts in the fielded military capabilities of states in the First World War timeframe.

7.4 Wartime Countermeasures Against Submarines

As explained in the previous chapter, pre-war countermeasures against submarines overlapped with countermeasures against torpedo attacks by torpedo boats and torpedo-armed destroyers. During the First World War, however, militaries encountered submarines in combat for the first time. As the number of operational submarines

increased and, more importantly, as Germany began to employ its submarines as commerce raiders, a role for which few had envisaged the employment of submarines prior to the war, militaries pursued new countermeasures, both technological and organizational-doctrinal to counter the threat posed by submarines to surface ships in general and merchant shipping in particular. It is important to recognize that these countermeasures did not need to be very effective let alone perfect to blunt if not negate the military effects of the latest advances in submarine technology on the fielded capabilities of states. Even imperfect countermeasures, and the latent threat posed by countermeasures capable of damaging or destroying a submarine, could impose virtual attrition by inducing caution on the part of submarine crews and interfering with submarine operations. A submarine force degraded in terms of effectiveness could, of course, itself impose virtual attrition on an enemy's surface fleet – inducing caution, incentivizing convoying, and so forth – and the effects of the wartime measure-countermeasure competition in undersea warfare were not therefore dichotomous. Even so, submarines had to pose a credible threat and be capable of sinking ships to be effective, and imposing virtual attrition on an enemy fleet was not the role that proponents of submarines before or during the war had in mind. The degree to which submarines could effectively sink ships is therefore an important metric and, as such, so is the degree to which countermeasures could blunt if not negate the threat posed by submarines to surface ships.

Given the limitations of First World War submarines in terms of speed, range, depth, armament, and sensing, the barriers to effective countermeasures were low particularly as enduring limitations in the area of propulsion technology forced

submarines to primarily operate on the surface. As a result, defending against submarines did not necessarily require developing effective means of damaging or destroying submarines. Rather, the primary objective was to interfere with submarine operations by deterring attacks and forcing submarines to dive when detected. With submerged submarines being much slower than surface ships and with torpedoes being unguided and limited in terms of range, surface ships had considerable scope for blunting the effectiveness of submarines. As a general rule, the fastest merchant ships, which were typically ocean liners, and the fastest surface warships, which included various types of cruisers as well as destroyers, were all but invulnerable to torpedo attack when moving anywhere close to their maximum speeds in open water other than in a chance engagement in which a ship moved directly toward an undetected (generally submerged) submarine. The largest and most heavily protected surface warships, namely heavy cruisers and battleships were also very difficult to sink – these large warships could sustain considerable damage – unless an unguided torpedo, which would hit a ship just below the waterline, detonated in a particularly vulnerable area and caused damage that the crew was unable to contain while the ship limped back to port for repairs. Submarines were, as such, most effective in (A) coastal waters in which surface warships were more restricted in their ability to maneuver or operate at high speed and (B) against merchant ships that were slow, unarmed, and unable to sustain damage. This is the operational and technological context in which militaries pursued anti-submarine countermeasures during the First World War.

7.5.1 Radio and Signals Intelligence

By 1914, radiotelegraphs were standard equipment in navies but were primarily restricted to larger vessels. As early as the Russo-Japanese War, militaries monitored enemy radio traffic and undertook efforts to secure their radio communications. The risks associated with not doing so were well recognized and were put on stark display on the Eastern Front in 1914 at the Battle of Tannenberg where German forces intercepted unencrypted Russian radio communications that offered information that contributed to a decisive German victory.[24] Given this backdrop and the inherent reliance of surface ships, particularly large groups of surface ships, on radio to effectively communicate with one another and with land-based headquarters, signals intelligence capabilities became increasingly central to naval operations, particularly in the North Sea. The opposing British and German fleets warily monitored one another's movements while anticipating an opportunity to catch their opponent by surprise and attain victory in a decisive engagement. As radio increasingly became standard equipment on submarines, militaries unsurprisingly applied their fast-growing signals intelligence capabilities to the purposes of countering adversary submarines.

Submarines had to operate on the surface to utilize radios and, for practical reasons, did so at night. In daylight, a submarine could be detected and forced to dive. The submarine would, therefore, be unable to receive radio messages. While operating in darkness, a submarine could send and receive information over radio without interruption. Darkness would not, however, conceal even encrypted radio transmission

[24] Frederick E. Jackson, "Tannenberg: The First Use of Signals Intelligence in Modern Warfare," *Strategic Research Project* (Carlisle, PA: US Army War College, 2002)

from signals intelligence capabilities, and the North Sea in particular was constantly monitored for radio traffic. Not only would intercepted radio transmissions give away the presence of a submarine and assist in determining its operating area, but militaries could also employ cryptanalysis to decode encrypted messages. The United Kingdom led the way in intercepting and decrypting naval radio traffic and adeptly used this information to route convoys (later in the war) and determine priority areas for surface warships to patrol. Signals intelligence was no panacea, but its very existence considerably undermined the capability gains offered to states through the application of advances in radio technology to submarines.

7.5.2 Aircraft, Whether Fixed-Wing or Lighter-Than-Air

Until the early 1910s, submarines, as well as all surface ships, could operate in complete disregard of the activity of the few fixed-wing aircraft in existence. Early fixed-wing aircraft were too limited in terms of range, speed, and payload to practically attack submarines and could not be equipped with a radiotelegraph through which to report the position of a submarine. By the start of the First World War, however, fixed-wing aircraft technology was characterized by rapid performance growth and the protracted nature of the war led states to make unprecedented resource allocations aimed at both dramatically increasing the number of fixed-wing aircraft available for use and developing progressively more capable aircraft suitable for a wider range of roles. During the war, the growing number of aircraft in service, including the subset of aircraft that were equipped with a radiotelegraph, were employed on primarily unarmed patrols over coastal waters. Although sensory acuity was limited to what the pilot and/or observer

could see beyond the cockpit, aircraft could quickly sweep an area, such as the approaches to a port, in a manner that surface vessels could not.

All things considered, fixed-wing aircraft did not offer states major gains in terms of anti-submarine warfare capabilities in the First World War timeframe but nevertheless contributed to an evolving context in which surfaced submarines and surface ships were increasingly likely to be detected in open water beyond the visual range of enemy warships. With submarines forced to primarily operate on the surface as a result of enduring limitations in propulsion technology, advances in aircraft technology helped to blunt the capability gains offered by larger longer-range submarines built during the war.

While often overshadowed by fixed-wing aircraft, lighter-than-air aircraft also played an important role in naval operations as these offered both Germany and the United Kingdom the ability to persistently monitor large areas of the North Sea in a manner that fixed-wing aircraft could not and in a context in which large surface ships risked rapid interception by enemy warships. Lighter-than-air aircraft were particularly effective in undertaking patrols aimed at detecting submarines operating on the surface given the greater practicality of employing radiotelegraphs on an airship relative to a small and underpowered fixed-wing aircraft with one or two crewmembers. Given the relatively low threat posed by air defenses over water, lighter-than-air aircraft could also be used to monitor one's own coastal waters and thereby detect surfaced submarines on the approaches to major ports. With radiotelegraphs, observers in lighter-than-air aircraft could notify shore stations of submarine activity and thereby make it more difficult for a submarine to approach a major port to lay in ambush without being detected.

7.5.3 Naval Mines

Undersea warfare tends to be inaccurately characterized as a competition between torpedo-armed submarines and surface ships. As explained previously, however, surface ships also employed torpedoes, and naval mines, the other underwater armament, were similarly employed by submarines against surface ships while simultaneously constituting a major and often underappreciated threat to submarine operations. While First World War submarines were incapable of submerging more for than several hours and therefore primarily operated on the surface, a submerged submarine was entirely oblivious of its surroundings and navigated with the aid of nautical charts, a compass, and a clock. A submarine could, as such, run into a minefield. Because naval mines are unsupervised weapons that cannot discern between friend or foe, a submarine could become a victim of a mine that it deployed or one deployed by friendly forces. In shallow, non-turbid water in high visibility atmospheric conditions during the daytime, as could be found along much of the coastline of the Mediterranean Sea, surfaced submarines and surface ships could visually detect many naval mines. In the northern Atlantic Ocean and the North Sea, however, mines were very difficult to detect, and this made mines effective against both surface ships and submarines.

Naval mines are essentially statistical weapons primarily intended to interfere with naval operations – many mines failed or became detached from their moorings, most never detonated, some were neutralized by enemy forces, and many that detonated did not inflict the desired level of damage. Even so, mines were a constant latent threat to both surface ships and submarines, and submarines were particularly susceptible to naval mines for two reasons. First, First World War submarines could not dive very deep and

had every incentive to operate at periscope depth while submerged given the absence of any means of underwater sensing, thereby making submarines susceptible to mines deployed at a depth required to target large surface ships. Second, First World War submarines could not operate very deep and were primarily limited to shallow coastal waters, particularly in areas such as the North Sea in which truly deep water is uncommon. As a result, submarines could not practically operate at greater depths to avoid most naval mines. Given these dynamics and the threat posed to surface ships by submarines, the approaches to ports, both civilian and military, were heavily mined – with the military deploying the mines maintaining records of the locations of the minefields and carving out channels through which its own ships could navigate. As a result, submarines could not operate freely in coastal waters and faced a major risk of running into naval mines.

7.5.4 Convoying

Although the aforementioned developments in anti-submarine countermeasures undermined the effectiveness of submarines, the single most consequential development in the First World War was the practice of convoying. For as long as warships were highly differentiated from lightly armed if not unarmed cargo ships, militaries employed some of the former to protect groups of the latter. By the First World War, militaries were familiar with convoying and exhibited variance in terms of their willingness to adopt the practice to protect their surface ships. The largest and most powerful surface warships, heavy cruisers, and battleships, invariably operated in groups and, no less importantly, were invariably escorted by a combination of light cruisers and destroyers to counter the

threat posed by torpedo boats, submarines, and torpedo-equipped fast destroyers. When it came to merchant shipping, however, privately owned ships were largely left to their own devices and tended to select their own routes and schedules.

In a context in which the primary threat to merchant ships in open water were cruisers as well as armed merchant ships and ocean liners serving as auxiliary cruisers, convoying was possible but simply put a lot of ships together that could be readily destroyed by gunfire from armed ships that unarmed merchant ships could neither outrun nor outfight. As a result, if convoys were to be formed to defend against cruisers and auxiliary cruisers, these would have to be escorted by similarly armed ships, which is to say other cruisers and auxiliary cruisers. For these reasons, the United Kingdom, the country with the largest merchant fleet to defend and the country with the greatest reliance on maritime trade, organized escorted convoys to transport the British Expeditionary Force to France at the start of the war but was otherwise reluctant to implement convoys for merchant ships let alone assign ships to escort convoys of merchant ships. Instead, the British military, both the navy and the army, aggressively set out to destroy Germany's capacity to undertake cruiser warfare on the high seas at the start of the war by deploying cruisers to neutralize German warships operating beyond the North Sea, capturing Germany's colonial ports (which could resupply German commerce raiders) and, not unimportantly, going so far as to sever German submarine telegraph cables worldwide and capture a long-range radio station in German Togoland along the Gulf of Guinea that allowed distant German commerce raiders to coordinate operations with Berlin.[25] Stated differently, the United Kingdom largely neutralized

[25] Hew Strachan, *The First World War in Africa* (Oxford University Press, 2004)

German ships on the high seas and bottled up the bulk of the German surface fleet in the North Sea at the start of the First World War, thereby securing the navigation of merchant shipping from (surface) commerce raiders outside the contested waters of the North Sea.

Although the German submarine force is widely associated with the targeting of merchant ships in the First World War, German submarine operations, like those of every other country, began with operations focused on attacking warships and undertaking coastal defense patrols. By the fall of 1914, German submarines began to target merchant ships, but this was initially restricted to the North Sea and therefore targeted shipping transiting between the British Isles and Scandinavia. Moreover, the ad hoc and largely unexpected employment of submarines as commerce raiders was initially undertaken under the terms of prize rules (cruiser rules) by which a submarine would warn a merchant ship of impending attack, accept its surrender, and allow the crew onto lifeboats before a boarding party scuttled the ship – torpedoes were a scarce commodity and could not be wasted on a stationary merchant ship, most of which (such as fishing boats) were in any event too small to justify the use of a precious torpedo. In February 1915, however, Germany declared the waters surrounding the British Isles to be a war zone (partially in response to the arming of some British merchant ships to resist German submarines operating under prize rules), thereby initiating the first of two fateful periods of unrestricted submarine warfare.

For as long as submarines operated according to prize rules, First World War submarines were not very effective in the sinking of non-stationary merchant shipping. This, to be clear, does not mean that submarines could not inflict shipping losses. Rather, it simply means that losses, specifically British losses, were limited and manageable and

that losses were largely concentrated in the North Sea. With the German submarine force being finite in number, with not all submarines being operational at any given time, and with not all submarines tasked with commerce raiding, the United Kingdom did not find itself in such a dire position as to urgently require convoying or assigning surface warships to the role of convoy escorts to protect against attacks from submarines. Unrestricted submarine warfare, however, changed this dynamic entirely, and, not unimportantly, this shift in German operational practice coincided with the entry into service of growing numbers of newly built larger, longer-range German submarines that could readily operate across British coastal waters. These submarines carried more torpedoes and were equipped with one or two medium caliber guns, thereby offering the armament required for a submarine to effectively operate in contravention of hitherto customary prize rules in the role of a commerce raider.

As noted earlier, the United Kingdom was familiar with the practice of convoying, but it did not adopt this practice to protect its merchant shipping following the start of unrestricted submarine warfare. As a result, the German submarine force was able to sink previously unimaginable numbers of British and neutral shipping for a time with submarines. A decade prior, not only were submarines not developed and deployed for the role of a commerce raider, but then extant submarines were wholly unsuited for the role even if prize rules were to be flouted. Partially as a result of this, and largely as a result of dynamics unique to the British navy, the United Kingdom was slow to recognize the efficacy of convoying as a countermeasure to submarine commerce raiders and was instead influenced by a variety of misconceptions concerning the opportunity cost in terms of shipping efficiencies lost, the number of escorts required, and the magnitude of

the threat posed by submarines to groups of merchant ships moving in a group when compared with merchant ships operating independently. As a result of these dynamics, the historical record is tainted by a formative experience in which a "new" military technology was highly effectual for a time – in a previously unanticipated role – not as a result of its monadic attributes – strengths – but as a result of the failure of a state – the United Kingdom – to adeptly respond with the appropriate available countermeasures.

The United Kingdom did not implement convoying for its merchant ships and foreign merchant ships transiting to and from British ports until May 1917, which is to say after Germany initiated a second period of unrestricted submarine warfare. Once it did so, the effectiveness of the German submarine force declined precipitously. Admittedly, convoys were implemented around the same time as the United States formally entered the conflict, a not insignificant confounding variable, but there is nevertheless a case to be made for the effectiveness of convoying on the basis of three considerations. First, there were no substantial advances in (fielded) late-war countermeasures against submarines that could distort the effects of convoying on measure-countermeasure dynamics. Second, while the entry of the United States into the conflict dramatically increased the supply of merchant ships that could offset losses to German submarines (including seized German merchant ships interned at hitherto neutral American ports), the German submarine force was at its zenith in 1917 in terms of the size and the number of the latest, more capable submarines ready for operational patrols. Third, there are a series of analytical reasons as to why convoying would be effective against submarines serving as commerce raiders independent of historical experience that the empirical record supports.

Unlike other wartime countermeasures discussed earlier in this chapter, convoys did not damage or destroy submarines or even contribute to the damage or destruction of submarines. Convoys did not, however, need to do so as First World War era submarines were poorly positioned to locate surface ships let alone sink merchant ships in the first place. Notwithstanding the many technical limitations of submarines, all submarines, and indeed all surface ships, encounter a major challenge in detecting, tracking, and intercepting enemy ships in open water (i.e., beyond the approaches to a port). Given the slower speed and, in the First World War context, inferior sensory acuity of submarines relative to surface warships, German submarines faced a profound open-water search problem. Simply put, there was a lot of water for a submarine with an exceedingly limited sensor range to cover in order to detect a moving surface ship. With merchant ships moving in open water, it was difficult for a relatively slow submarine (even whilst surfaced) to intercept a detected merchant ship or group of merchant ships so as to employ inaccurate unguided torpedoes of limited range.

By implementing convoying, the United Kingdom, in effect, dramatically reduced the average density of merchant ships in open water per unit of time. While convoys could encounter – run into – German submarines, there were too few German submarines operating across a vast expanse of water for this to realistically happen more than once or twice for a given convoy, no German submarine carried enough torpedoes to sink an entire convoy of merchant ships, and the combination of armed merchant ships and warships serving as escorts could deter German submarines from attempting to pursue a convoy and/or attack the convoy with deck guns (for which there was a much greater ammunition supply onboard). With individual German submarines less likely to

encounter merchant ships without benefitting from an increase in endurance (i.e., increase in patrol duration), German submarines therefore increasingly returned to port without being able to sink ships with deck guns let alone torpedoes. As a result, convoying both compounded the German submarine forces' open-water search problem and imposed virtual attrition, thereby dramatically undermining the effectiveness of submarines as merchant raiders even if convoy escorts did not sink a single submarine. Convoying, therefore, demonstrated that the nominal capability gains in the area of commerce raiding offered by advances in submarine technology could be readily blunted if not negated with appropriate organizational-doctrinal countermeasures.

While no single countermeasure toward submarines was individually effective, the collective of the aforementioned countermeasures greatly interfered with submarine operations and thereby blunted the effects of advances in submarine technology on the fielded military capabilities of states. Submarines were forced to primarily operate on the surface and were extremely vulnerable when doing so. As a result of enduring limitations in the area of propulsion technology, First World War submarines could only briefly operate submerged, and while this increased the survivability of a submarine for a time, a submerged submarine was a largely ineffectual submarine given the absence of practical means of underwater sensing. To employ the small number of torpedoes onboard, First World War submarines had to either operate at periscope depth or on the surface. Every countermeasure that incentivized if not forced a submarine to dive allowed surface ships to evade an attack and, over repeated instances of diving and then resurfacing to recharge batteries, depleted the endurance of a submarine such that it would have to return to port. Over multiple operational patrols and many months, this could result in a situation in

which submarines could sink some ships but never so many or at such a pace that, in a protracted conflict, shipping losses would not be offset by new construction or, no less importantly, by losses of submarines. As a result of these dynamics, countermeasures greatly blunted the capability gains that advances in submarine technology offered states in the First World War timeframe.

Chapter 8: Submarine Technology ca. 1918-1939

The First World War experience with submarines offers limited evidence in support of the proposition that advances in submarine technology engendered sizeable and abrupt shifts (gains) in the fielded military capabilities of adopting states. Submarines were employed in a protracted high-intensity conflict before submarine designs were matured and before sizeable numbers of the latest, more capable designs were in service. Although submarines sank considerable numbers of primarily unarmed merchant ships, submarines proved far less effective against warships and merchant ships protected by escorts and merchant ships operating in convoys. The monadic limitations of submarines and the susceptibility of submarines to a wide range of countermeasures considerably blunted if not negated the military effects of advances in submarine technology on the fielded military capabilities of states. While this weakens support for the proposition that technological change engenders sizeable and abrupt shifts in the fielded military capabilities of states, it is possible that submarine technology was prematurely employed in combat and that, over a longer timeframe, we may find greater evidence of technological change engendering sizeable and abrupt shifts in the fielded military capabilities of states. By extending the longitudinal analysis into the Interwar Period, we can assess whether an additional two decades or so of technological development and submarine construction resulted in a situation in which advances in submarine technology engendered more significant effects on fielded military capabilities.

Although advances in submarine technology are a recurring area of focus for scholarship examining the effects of advances in military technology, the Interwar Period tends to receive far less attention than the First World War, Second World War, or the

Cold War when it comes to submarine technology. This reflects the tendency for scholarship to focus on wartime cases that offer a large number of observables and cases in which a technological breakthrough can be identified. The Interwar Period experience does not offer readily accessible observables in the form of combat performance and did not encompass any substantial breakthrough in submarine technology. Instead, the Interwar Period was one in which submarine technology became progressively more refined and, in the absence of a technological breakthrough, approached a state of maturity, thereby creating the conditions in which submarines and anti-submarine capabilities were employed on an even greater scale in the Second World War. The submarines employed during this later conflict did not, of course, emerge in a vacuum and were, by and large, designs that were developed, tested, and, for the most part, both first manufactured or first deployed during the Interwar Period. As such, it is analytically productive to continue the longitudinal analysis of the military effects of advances in submarine technology into the Interwar Period to determine whether there is evidence of technological change engendering sizeable and abrupt shifts in the fielded military capabilities of states.

Submarine technology advanced consistently if gradually throughout the Interwar Period but the implications in terms of shifts in fielded military capabilities were limited. In large part, this is because submarine technology began to plateau from the levels reached toward the end of the First World War in the absence of a fundamental breakthrough in the area of propulsion technology. To be clear, this is not to say that there was no technological progress during the Interwar Period, but only that the substantive capability gains engendered by new areas of performance growth were

limited by the absence of a breakthrough in propulsion technology. Simply put, Interwar Period submarines remained submersibles that primarily operated on the surface. As such, advances in submarine technology over the Interwar Period are best characterized as gradual refinement toward a state of maturity.

The First World War offered Interwar Period militaries and the designers of submarines wartime experiences to reflect upon, something that neither could benefit from in the lead-up to 1914. At first, militaries pursued marginally upgraded versions of late-war designs but, over the Interwar Period, militaries increasingly pursued improvements, and while novel concepts were experimented with, resources were primarily directed toward refining existing technology and building and fielding successive modestly improved designs. As such and with the notable exception of the performance parameter of maximum operating depth, submarines built in the 1920s tended to be broadly comparable to submarines built in the late 1910s while submarines built throughout most of the 1930s tended to be broadly comparable to submarines built in the 1920s. To appreciate what this dynamic meant in practice, I will examine in turn developments in the areas of armament, sensing, communication, and propulsion.

8.1 Advances in Submarine Armament

During the Interwar Period, only the smallest "coastal" submarines, which generally displaced no more than 300-400 tonnes, were designed to exclusively employ torpedoes; all other submarines were explicitly conceived as gun- and torpedo-equipped vessels even when existing in blueprint form. This reflected upon wartime experiences that highlighted the importance of a secondary armament with a much larger ammunition

supply and the reality that submarines could only submerge for little more than a day (without surfacing to recharge their batteries) and were as such submersibles primarily operating on the surface. Given the considerable size and weight of torpedoes, even the largest submarines could carry little more than a dozen (larger diameter, larger warhead, longer-range) torpedoes and could launch even fewer at a time (i.e., there were fewer torpedo tubes than onboard torpedoes; the remaining torpedoes were reloads). In contrast, a large ocean-going submarine, which generally entailed a displacement of over 700 tonnes, could readily carry over a hundred rounds of ammunition for a medium caliber gun (generally in the range of 70-100 mm in bore diameter). While submarines were poor gunnery platforms and although small salvos fired by one or two guns were inaccurate against moving targets other than at close ranges (with flat trajectory fire) at which a submarine would itself be extremely vulnerable to gunfire, Interwar Period submarines could only otherwise employ unguided torpedoes that were little changed from those available to militaries at the end of the First World War.

As explained earlier, while the introduction of the Whitehead torpedo in the 1870s constituted a technological milestone, its power source, a compressed air engine, offered poor performance in terms of range and speed. By the early twentieth century, militaries developed torpedoes equipped with (dry) heater engines and, in time, wet-heater engines, both of which entailed the use of kerosene in a combustion process. While Interwar period torpedoes became increasingly refined on the basis of wartime experience – but not necessarily increasingly reliable – torpedoes were still unguided, and it remained difficult to hit moving targets. Compounding matters, surface warships experienced considerable performance growth in terms of maximum speed over the Interwar Period

from around twenty knots to around thirty knots. This compounded the challenges of fire control for submarines and drove militaries to pursue torpedoes that offered greater speed-range, with a torpedo operating at a faster speed setting having a shorter range and vice versa.

Building on war-tested wet-heater torpedo engine technology, militaries experimented with increasing the supply of onboard oxidizer throughout the Interwar Period. Compressed air, which wet-heater torpedoes used in a combustion process, has a relatively low oxygen content, being approximately 21% oxygen and 78% inert nitrogen, and designers therefore had a clear incentive to increase oxygen levels in the compressed air tank contained within a torpedo. Doing so, however, came at the cost of turning the torpedo, itself an intrinsically volatile combination of a high explosive warhead and combustible fuel stored store within the hull of a submarine, into a major safety hazard. While multiple countries pursued advances in torpedo propulsion technology, most efforts did not result in substantially improved new torpedoes entering service. The United Kingdom, for its part, developed and deployed torpedoes that used oxygen-enriched compressed air, and, separately, the country also developed diesel-type burner engines, but the performance gains offered were relatively modest. By the mid-1930s, however, Japan pioneered torpedoes that made use of compressed oxygen and did so despite the known risks (the consequences of this acceptance of risk were largely invisible in peacetime but became highly taxing during the Second World War).[26] By multiplying the supply of oxidizer available to the onboard engine by a factor of approximately five, Japan's Type 93 ("Long Lance") and Type 95 torpedoes – employed

[26] Consider, for example, the discussion of American intelligence on Japanese torpedo development in Thomas Mahnken, *Uncovering Ways of War: U.S. Intelligence and Foreign Military Innovation, 1918–1941* (Ithaca, NY: Cornell University Press, 2002).

by surface ships and submarines, respectively – offered considerably greater speed-range and, in principle, offered Japanese submarines torpedoes with ranges comparable to the practical range of the latest naval guns. As such, Japan benefitted from unmatched nominal gains in this narrow area of its fielded military capabilities by the start of the Second World War.

Concurrent with these developments in the United Kingdom and Japan, Germany discretely developed and later deployed electrically powered torpedoes employing lead-acid batteries and an electric motor instead of a combustion engine. The resulting torpedo, the G7e, which was deployed in the late 1930s, was simpler and less expensive to manufacture than wet-heater torpedoes and, not unimportantly, did not generate air bubbles from a combustion process and did not therefore generate a tell-tale visual signature of an active torpedo. Given the lower energy density and specific energy of electric batteries compared to a petroleum-based fuel such as kerosene, however, the G7e was inferior in terms of speed-range and electrically powered torpedo did not therefore supplant wet-heater torpedoes in German service.

Given the aforementioned trade-offs, there is a case to be made that although torpedo technology improved in terms of speed-range – less so reliability – over the Interwar Period, performance growth was relatively modest in terms of the practical effects on fielded military capabilities, and major performance gains required significant changes to propulsion technology that came with significant downsides relative to well-established wet-heater torpedo technology. As such, militaries continued to primarily employ torpedoes of the decades-old wet-heater configuration over the Interwar Period, and these were, all things considered, little better than the torpedoes employed by the end

of the First World War. Moreover, all torpedoes deployed during the Interwar Period, including the latest designs such as the Japanese Type 93 and the German G7e, were unguided "straight runners." As a result, all torpedoes were only as accurate as the (monadic) fire control solution computed by a submarine's crew, and moving ships, particularly maneuvering ships, therefore remained difficult targets to hit with a single torpedo over long distances, thereby necessitating the launch of a salvo, or spread, of torpedoes, preferably at close range to attain a higher probability of impact. With a very limited supply of torpedoes onboard even the largest of Interwar Period submarines, a submarine-launched torpedo remained a greater threat to a surface ship in principle than it was in practice, and advances in torpedo technology did not result in a substantive disjuncture from earlier dynamics covered in the previous chapters.

8.2 Advances in Submarine Sensing Technology

With torpedoes being unguided, a substantial advance in the area of sensing was required to significantly improve the capabilities offered by the latest submarine designs. By 1918, early forms of both active sonar (i.e., ASDIC, to use then contemporary British terminology) and passive hydrophones were developed by the United Kingdom and the United States. Over the Interwar Period, both types of underwater sensors were increasingly deployed and subject to further technological development. While the implications of these technological advances for the military capabilities offered by the latest advances in submarine technology are best understood in terms of measure-countermeasure dynamics and will as such be examined later in this chapter, it is

important to also consider what active sonar and passive hydrophones offered a submarine and its crew.

In principle, active sonar allowed a submarine to independently locate targets even if the target vessel was stationary and not generating hydrodynamic or propulsion noise. In practice, however, active sonar was not well suited for detecting surface ships, was very limited in terms of range and field of view, and could only be employed against surface ships while a submarine was submerged (at a time when propulsion technology forced submarines to primarily operate on the surface). Additionally, employing active sonar would give away the presence of a submarine to any submarine or warship equipped with passive hydrophones, which were increasingly common over the Interwar Period. For these reasons, active sonar was employed on surface ships, not submarines, during the Interwar Period and, if employed on a submarine, would be more useful as an acoustic rangefinder to assist in the computation of a firing solution for unguided torpedoes than a means of detecting enemy ships.

Passive hydrophones, which were also first developed toward the end of the First World War, constituted a far more practical means of sensing but a submarine operating on the surface was far from an optimal platform for deploying a passive hydrophone. Hydrodynamic noise and noise from the submarine's own diesel propulsion system, which powered a surfaced submarine, made it difficult to detect other vessels. At the same time, passive hydrophones offered submarines a means of sensing once submerged. This allowed submarine crews to operate at greater depths (i.e., below periscope depth), which otherwise posed navigational challenges and left a submarine blind to its surroundings, and this, in turn, incentivized submarine designers to develop hulls capable

of withstanding the increases in hydrostatic pressure encountered at greater depths. As such, passive hydrophones had an arguably greater effect on how submarines could operate – such as diving at greater depths to evade depth charges – than as sensors for detecting enemy shipping in the vastness of open waters.

Given that Interwar Period submarines primarily operated on the surface and relied on human eyes assisted by telescopic optical devices – radar was not available until the late 1930s and the first radars deployed on ships and on land were unsuitable for use on submarines, the practical effects of advances in underwater sensing were limited. When submarines briefly submerged – invariably for no more than twenty-four to forty-eight hours unless the crew was willing to risk fully depleting the battery – a submarine could use passive hydrophones to be aware of proximate shipping activity but would invariably have to operate at periscope depth and rely on human eyes assisted by optical devices to detect, identify, track, and engage enemy vessels. As a result, Interwar period submarines equipped with passive hydrophones were essentially as limited in terms of sensory acuity as late First World War submarine designs.

8.3 Advances in Communications Technology

While Interwar Period submarines remained submersibles with poor sensory acuity and a small load of faster and longer reaching but nevertheless unguided torpedoes, one area of technological development largely unrelated to naval let alone submarine technology considerably expanded the capabilities that submarines offered states. As explained earlier, radio technology, specifically radiotelegraphy, existed before the First World War but was primarily employed by larger surface warships. During the First World War, a

wider range of naval vessels, including submarines, were equipped with radio sets, and increasingly sophisticated and reliable radios were employed throughout the Interwar Period on the heels of advances in radio technology primarily made by civilians for civilian purposes.

With more effective and more reliable radios available, submarines could better coordinate operations with surface ships, land-based headquarters, and, no less importantly, other submarines. This allowed submarines to transmit and receive the reported locations of enemy shipping and, in terms of inter-submarine communications (whether direct and/or through intermediaries at land-based headquarters), coordinate attacks on groups of surface warships and convoys of merchant shipping. Radio technology therefore allowed submarines to share what amounted to sensor data and thereby mitigate the profound search challenge faced by individual submarines in open water. Advances in radio technology allowed submarines to operate in groups – to operate in so-called wolfpacks – that could coordinate with one another to intercept a detected force of warships or a convoy of merchant at one or more locations in what could be multi-day-long engagements involving dozens of ships and large groups of submarines (primarily transiting on the surface to keep up with their intended targets). Stated differently, radio communications allowed militaries to operate submarines in groups and thereby partially overcome the limitations of individual submarines in terms of sensory acuity and munitions payload. This was not a new dynamic in the 1930s – the earliest (coastal) submarines were, after all, intended to operate in coastal defense roles in flotillas composed of multiple submarines in much the same manner as torpedo boats (albeit without radios in the lead-up to the First World War). A flotilla of coastal

submarines or torpedo boats operating in coastal waters was, however, restricted to operating within several dozen kilometers of the coastline over no more than several days. With radio communications, dozens of much larger submarines could coordinate operations hundreds if not thousands of kilometers from their home ports over several weeks.

Although advances in radio technology improved the effectiveness of submarines over the Interwar Period, it is important to recognize that because radios were not exclusive to submarines, advances in radio technology also increased the threats faced by submarines. Radio emissions could, after all, be detected, thereby giving away the position of a submarine, and even if communications were encrypted, an adept adversary could endeavor to decrypt the messages, thereby giving the adversary insight into a submarine's missions and its past, present, and future locations. During the First World War, all leading militaries undertook extensive signals intelligence efforts in support of naval operations, and this measure-countermeasure competition in intelligence gathering continued throughout the Interwar Period. Compounding matters, advances in radio technology also meant that once detected through some means, a submarine could expect essentially every merchant and naval ship in the area to be aware of its presence and some of the alerted ships could actively search for the submarine, thereby forcing it to draw upon its finite battery capacity to submerge – rendering it incapable of sending or receiving messages over radio – so as to avoid detection with the net result of interfering with submarine operations in a manner that blunted whatever capability gains that advances in radio technology nominally offered submarines.

8.4 Advances in Propulsion Technology

Whatever the armaments, sensors, and communications systems employed on the latest submarines, the performance of a new submarine and, as such, the capability gains it offered states, was to a large extent shaped by the design of a submarine's hull and its propulsion system. Stated differently, naval architecture and propulsion technology, not armament, sensor, or communications technology, determined how fast and how far a submarine could travel while surfaced and submerged, respectively, and how much weight of volume could be allocated to armaments, sensors, and communications systems. As a result, naval architecture and propulsion technology determined the roles for which a new submarine would be effective, and the capability gains it offered states relative to prior submarine designs.

As noted at the start of this chapter, Interwar Period submarines were, by and large, more refined versions and derivates of late First World War submarines. By the 1930s, successive gradual advances increased the performance gap with earlier submarines but leaving aside the issues of reliability and refinement, submarine technology primarily improved in terms of the maximum depth at which a submarine could operate. As a result of advances in metallurgy and manufacturing techniques such as welding, Interwar submarines could operate at progressively greater depths that, by the late 1930s, tended to be 100-200 meters for the largest ocean-going submarines. By operating at greater depths, submarine crews could eliminate the possibility of colliding with the submerged sections of a ship's hull and fishing nets, thereby allowing a submerged submarine to safely operate in busy shipping lanes as bathymetric conditions allowed.

While advances in metallurgy and manufacturing techniques allowed militaries to operate sturdier, safer, and more reliable submarines, the capability gains were limited, and it is important to appreciate why this was the case before proceeding to examine developments in the area of propulsion technology. In principle, greater operating depth allowed submarine crews to dive deeper to reduce the probability of damage or destruction from an unguided depth charge and, to a lesser degree, evade certain types of naval mines (such as moored mines maintained at a specific depth). In practice, however, greater operating depth had limited applications on a day-to-day basis because deep water is typically found in open waters. The enduring limitations of battery technology meant that submerged submarines were, however, poorly positioned to detect, track, and intercept surface ships, whether naval ships or merchant ships, traveling at cruise speed which surface ships tend to do in open waters devoid of hazards to navigation, unlike coastal waters. As a result, submarines primarily operated on the surface as submersibles in open water, thereby rendering greater operating depth moot in many, though not all, circumstances.

In some politically salient military-geographical contexts, such as the North Sea, bathymetric realities meant that submarines could not generally dive beyond fifty meters in depth without grounding or risk colliding with an undetected or uncharted shipwreck. In other areas, such as the open waters of the Atlantic and Pacific oceans and parts of the Mediterranean, a greater operating depth gave submarine crews a considerable safety margin that was not available in earlier submarines that risked leaks and the catastrophic implosion of the pressure hull when operating at depths of fifty or more meters. Even so, no submarine in existence during the Interwar Period could safely operate at depths much

greater than 200 meters and this meant that submarines operating in truly deep areas of open water would risk the catastrophic implosion of the pressure hull long before the submarine would reach the sea floor. As a result, improvements to the reliability of onboard machinery, above all the electric propulsion system that powered a submerged submarine, were arguably of far greater practical importance than increases in hull strength as far as submerged performance was concerned.

As explained in previous chapters, human-powered submarines were a possibility as early as circa 1800, and self-propelled submarines were a possibility as early as circa 1860 but submarines were impractical, ineffective, and, not least, uncompetitive means of employing self-propelled torpedoes until the late 1890s. For the most part, this was a result of the unavailability of propulsion systems that were suitable for both surfaced and submerged locomotion. By the early twentieth century, however, leading militaries converged upon the use of an internal combustion engine (later exclusively of the Diesel-cycle variety) alongside the combination of an electric motor and an electric battery with the latter being the sole means of underwater propulsion.

This hybrid propulsion arrangement served militaries well in the First World War but enduring limitations in battery technology, namely the poor energy density and specific energy of batteries when compared with petroleum products, meant that submarines could not travel very fast or for very long while submerged. As a result, the limitations of propulsion technology dictated that submarines would have to primarily operate on the surface and this unsurprisingly led to submarines with hull designs optimized for operations on the surface, thereby resulting in the near-universal reality that essentially every submarine built before the 1950s had a higher maximum speed while

surfaced (while powered by an internal combustion engine) than submerged (while powered by electric batteries). This dynamic continued throughout the Interwar Period as there was no substantial breakthrough in propulsion technology and no viable alternative to the employment of an internal combustion engine paired with the combination of an electric motor and electric batteries. As a result, Interwar period submarines remained submersibles that primarily operated on the surface and only briefly submerged to approach or escape an enemy ship or otherwise evade detection.

Although there was no substantial breakthrough in the area of submarine propulsion technology, it is important to recognize that this is not to say that there were no technological advances; only that the gains engendered in fielded military capabilities were limited and primarily manifested in terms of the increased mechanical reliability and power output of the diesel engines used in submarines (and, to a lesser extent, advances in battery technology). Even so, three areas of technological development are worth briefly highlighting.

First, the United States led the way in developing and deploying true diesel-electric propulsion systems in which the diesel engine was not mechanically connected to the propeller(s).[27] Prior to the introduction of such systems, all submarines employed a diesel engine directly connected to both the shaft(s) driving the propeller(s) and an electrical generator whereas a true diesel-electric configuration meant that the diesel engine would always power an electric generator – convert mechanical energy into electrical energy – and an electric motor would in turn drive the shaft connected to the propeller. Among other benefits, a true diesel-electric propulsion system allowed for the

[27] See Norman Friedman, *U.S. Submarines Through 1945: An Illustrated Design History* (Annapolis, MD: Naval Institute Press, 1995)

operation of the diesel engine at an optimal speed without dictating the speed at which a submarine itself had to operate. From the 1950s onward, true diesel-electric propulsion systems became universal on (non-nuclear) submarines, but these were not deployed outside the United States and, to a lesser degree, Sweden, during the Interwar Period.

The second and third notable areas of Interwar technology development respectively concern (A) snorkels through which a diesel-powered submarine could operate just beneath the water's surface to recharge its batteries and (B) efforts toward the development of air-independent propulsion systems that could allow a submarine to operate an engine (i.e., not only its batteries) for underwater locomotion without drawing in oxygen from the atmosphere. Snorkels were pioneered by the Netherlands and were only deployed on an essentially inconsequential scale during the Interwar Period and are therefore best examined in the context of the introduction of large numbers of snorkel-equipped submarines by Germany during the Second World War, a topic that will be covered in the following chapter. Air-independent propulsion systems were similarly not deployed during the Interwar Period and are also best examined in the context of the radically different German Type XXI and Type XXVII submarines developed late in the Second World War that will also be examined in the next chapter.

Although submarine technology gradually advanced during the Interwar Period, limitations in terms of armament, sensing, radio, and, most importantly, propulsion technology, restricted the magnitude of capability gains engendered to adopting militaries. Given that the gap between the latest submarines employed by leading militaries in politically salient state dyads was limited throughout the Interwar Period with the singular exception of Japanese torpedoes employing compressed oxygen, no

state developed and deployed an advance in submarine technology that engendered a major capability gain. Successively more refined submarine designs were just that and improvements in metallurgy, hull design, engine reliability, etc. did not offer states submarines with either substantially improved capabilities or submarines that were unaffected by enduring technological limitations.

8.5 Experimentation and Conceptual Innovation in Submarine Design

Although submarine technology largely plateaued during the Interwar Period as a result of enduring technological limitations, it is important to recognize that this did not reflect upon a lack of ambition on the part of leading militaries. Constrained though leading militaries were in the context of naval treaties and, for much of the Interwar Period, limited peacetime funding for the development and construction of new designs, the Interwar Period was characterized by considerable conceptual innovation in submarine technology. To be clear, experimental submarine designs developed, built, and tested during the Interwar Period were, by and large, unsuccessful and, as such, have not entered the canon of retrospectively effectual technological innovations that are the primary focus of scholarship on the military effects of technological change. Nevertheless, it is analytically productive to examine some of the more radical, and in that sense more innovative, concepts pursued in the Interwar Period as this offers an opportunity to develop a more holistic understanding of the enduring limitations of submarine technology and, as such, an opportunity to appreciate the limited military effects of technological advances during the Interwar Period.

By the end of the First World War, militaries had ample evidence that the combination of self-propelled submarines and self-propelled torpedoes did not herald the demise of surface warships. It was difficult for a submarine to successfully detect and engage let alone damage or sink a moving surface warship in most circumstances and submarines were exceedingly vulnerable to countermeasures. While no major military willingly abandoned submarine technology – Germany did so unwillingly for a time and covertly pursued advances in submarine technology despite legal restrictions on doing so – following the end of the war, early post-war submarines were not suitable for the roles and military-geographical requirements of several leading militaries and these countries led the way in developing new conceptually innovative types of submarines.

8.5.1 Fleet Submarines

The British navy was one of the earliest adopters of self-propelled submarines but nevertheless found it difficult to productively employ submarines in its force structure. The earliest self-propelled submarines in British service were essentially small submersible torpedo boats restricted to coastal operations that offered little in the way of gains in fielded military capabilities given the preponderant state of the pre-1914 British navy. While larger submarines were developed prior to 1914 with the range and endurance required to operate throughout the North Sea, among other areas, these were too slow to operate alongside British surface warships. As a result, British submarines primarily operated independently of the British surface fleet during the war. Given the dearth of German merchant ships to sink in wartime, British submarines also necessarily

undertook very different missions than the submarines of other states including anti-submarine warfare against the German submarine force.

From the early 1910s, the British navy pursued a much faster "fleet submarine" that could keep up with surface warships and thereby operate in support of the surface fleet. Ideally, the thinking went, such a submarine could augment the vital scouting and screening roles fulfilled by fast torpedo-armed light cruisers and destroyers. By 1915, the British navy determined that even the latest diesel engines were inadequate to propel a surfaced submarine to a speed of twenty or more knots, which was required to operate alongside the surface fleet while it was moving at full speed (heavily armed and armored battleships were among the slowest warships in the surface fleet at the time). For this reason, the British navy turned to steam turbines for the new 1800-tonne displacement K-class submarines that entered service from 1917 through 1921. The British navy had a very difficult experience with the seventeen-strong K-class, with six submarines lost outside of combat operations. In many respects, this was not entirely surprising and for the most part, the shortcomings of the design could be attributed to the use of oil-burning steam turbines which, among other things, resulted in a slow-to-dive submarine given the need to secure the funnel before diving. Designers had considered steam propulsion for submarines since the 1880s and a small number of unsuccessful steam-powered submarines were built before the turn of the century. The British navy was, as such, aware of the limitations of steam propulsion in a submarine and pursued the K-class design because then existing steam turbines constituted the highest power output propulsion systems available at the time by a very wide margin.

While the K-class was, all things considered, a costly and expensive failure, the British navy did not abandon the concept of a high-speed fleet submarine until the early 1930s following the construction of three River-class submarines which made use of the latest in supercharged diesel engine technology but could nevertheless not meet requirements. Not only were designers struggling to develop submarines that could sustain high surface speeds while retaining desirable specifications required to be effective as a submerged submarine, but the latest surface warships were becoming faster such that even battleships were approaching 30 knots by the late 1930s and fast destroyers, which could also screen a surface fleet, could exceed 30 knots.

The United States also pursued fleet submarines in the Interwar Period, but variant military-geographical contexts resulted in different requirements than those pursued by the United Kingdom. In the United States, the objective was to develop a submarine that could operate in support of the surface fleet rather than alongside it. Driven by planning for a primarily naval war against Japan in the vast Pacific Ocean, the United States sought submarines that could undertake extremely long-range patrols in the Pacific Ocean to screen the American surface fleet and impose attrition on Japanese surface warships expected to head east before a (mutually) anticipated major decisive battles took place between the opposing surface fleets in the central Pacific.[28] Like the United Kingdom, the United States also aimed for a maximum surface speed of over 20 knots, but increasing range was far more important for the latter particularly as the latest large surface warships including battleships progressively exceeded the speed that could be practically achieved by a surfaced submarine.

[28] Edward S. Miller, *War Plan Orange: The U.S. Strategy to Defeat Japan, 1897-1945* (Annapolis, MD: Naval Institute Press, 1991); David Evans and Mark Peattie, *Kaigun: Strategy, Tactics, and Technology in the Imperial Japanese Navy, 1887-1941* (Annapolis, MD: Naval Institute Press, 1997)

Japan, for its part, also pursued fleet submarines for its navy, and, like the United States, it placed a heavy emphasis on range and endurance given its preparation for a potential conflict in the vast expanse of the Pacific Ocean. Unlike the United States, however, Japan increasingly focused on aircraft-carrying fleet submarines over the Interwar Period, and these will be separately examined later in this chapter.

All things considered, the fleet submarine concept was a geographical adaptation and not a reflection of any significant breakthrough in submarine technology. Sweden, for example, could have developed the most advanced submarine in the world to primarily operate in the Baltic Sea and around the Danish Straits and such a submarine would have been all but useless to the United Kingdom, United States, and Japan given the variant military-geographical contexts of these countries. At the same time, the latest fleet submarines developed and deployed by these countries would have likely been ineffective in the hands of Sweden in the Baltic. Stated differently, a larger submarine was not necessarily a better submarine unless a state placed tremendous emphasis on endurance and range and was willing to pay for it in terms of expense and design trade-offs. A larger submarine could, of course, operate at greater distances (whilst surfaced), carry more supplies, and carry more armament (including additional torpedo reloads) but it was not necessarily faster (by virtue of being larger), more maneuverable, faster to dive, etc., or, as such, a better submarine. Given the state of battery technology and notwithstanding the existence of a commensurably greater number of battery cells on larger fleet submarines, a larger submarine also did not inherently offer increased underwater speed or underwater endurance. At the same time, a larger submarine was more expensive and complex to build and operate, required a larger crew, and, despite its

greater size, was no less vulnerable to anti-submarine countermeasures. For these reasons, it is important to recognize that fleet submarines were developed *despite* enduring limitations in submarine technology and not on the basis of any breakthrough technological advances. If adjusting for size and displacement, there was little in a fleet submarine that was distinguishable from that which could be found in much smaller submarines and the emergence of long-range fleet submarines did not therefore engender particularly sizeable gains in the fielded military capabilities of states.

8.5.2 Cruiser Submarines

Whereas fleet submarines were large and fast (whilst surfaced) submarines intended to operate in support of surface warships, cruiser submarines were intended to operate independently as naval combatants and serve as submersible substitutes for large ocean-going cruisers on long-range deployments. Given the variant experiences with and requirements for submarines in the First World War, it is unsurprising that Germany pioneered the concept of a cruiser submarine (as a solitary long-range commerce raider), but it nevertheless built and operated very few submarines worthy of the designation by the end of the war and the country was prohibited from operating any submarines by post-war settlements. Cruiser submarines were, therefore, pursued during the Interwar Period by states that did not share Germany's acute (wartime) need for a long-range submersible that could substitute for (surface) cruisers as a commerce raider.

Like fleet submarines, cruiser submarines obtained their long-range by greatly increasing the size of the submarine to increase the amount of stored fuel and supplies. These large submarines typically displaced 2,000 tonnes or more (most First World War

submarines had a surface displacement in the region of ~400-1,000 tonnes) and notably complemented torpedoes with large caliber naval guns of over 100 mm bore diameter (fleet submarines, in contrast, were primarily intended to target surface warships against which a gunnery duel was essentially suicidal). Equipped with such armament and a sizeable ammunition store, cruiser submarines could, in principle, serve as long-range commerce raiders primarily operating on the surface while retaining the capacity to submerge to approach or escape an enemy ship or otherwise evade detection.

Whereas First World War submarines employed guns mounted on their decks as secondary armament out of necessity – there were too few torpedoes on board and all torpedoes were unguided – cruiser submarines were designed from the outset to employ one or more large caliber guns. The most ambitious and the most extreme manifestation of the cruiser submarine concept was developed by France. The singular *Surcouf* was the largest submarine built at the time by a wide margin and displaced over 3,200 tonnes whilst surfaced. The French submarine's superstructure was dominated by a pair of 203 mm guns in a highly restrictive turret arrangement built into the submarine's superstructure – it did not have an actual deck gun – and this gave the *Surcouf* submarine the firepower of a surface cruiser. On paper, the *Surcouf* was a formidable submarine and a major threat to merchant ships on the high seas.

While optimized for different roles, cruiser submarines had the same limitations as fleet submarines as the larger size was not accompanied by any technological breakthroughs. A cruiser submarine was consequently just as limited in terms of sensory acuity as any long-range ocean-going submarine and was just as vulnerable to anti-submarine warfare capabilities. Moreover, the capability gains nominally offered by

cruiser submarines were primarily endowed by large gun(s), but all submarines were inherently poor gunnery platforms. Low-elevation optical devices and rangefinders limited the range at which a cruiser submarine could accurately employ its guns, a submarine was not a stable firing platform outside of calm waters and this negatively affected the accuracy of gunfire, guns were limited in terms of elevation and traverse in the absence of a meaningful turret arrangement, and there was a limited supply of ammunition – and a limited number of barrels ready to fire – onboard to compensate for these limitations. Furthermore, if a cruiser submarine found itself in a gunnery duel with a surface warship or an armed merchant ship, the unarmored submarine was far more vulnerable to the effects of naval gunfire than its targets. Whereas a surface ship could expect to sustain some level of damage and remain afloat and operable, a submarine that received damage from gunfire risked the catastrophic failure of its pressure hull when it eventually submerged even as submerging was the best way to extricate itself from the situation. At the same time, a cruiser submarine's crew faced incompatible tactical realities given that the best position, heading, and speed to employ a large-caliber gun with limited elevation and traverse against a moving target was rarely the best position, heading, and speed (let alone depth) to employ torpedoes and vice versa. All things considered, cruiser submarines had the same limitations as fleet submarines – these were only larger submarine designs but did not benefit from any fundamental breakthrough in submarine technology and were no more capable or less limited than their much smaller counterparts in performance parameters other than range, endurance, and the firepower derived from larger-caliber gun-type armament.

8.5.3 Aircraft Carrying Submarines

The concepts of a fleet submarine and a cruiser submarine respectively emphasized speed and range, and range and gun-type armament and were respectively intended to operate in support of and independent of formations of surface warships. While several militaries experimented with operating aircraft from submarines over the Interwar Period, only Japan built and fielded large numbers of such submarines to serve in a scouting role for its surface fleet.[29] Japanese aircraft carrying submarines had facilities to store, launch, and recover one or two floatplanes – some designs were equipped with catapults while others used a crane to lower the stored floatplane into the water – and were, as such, very large submarines.

Given Japan's military-geographical context, these submarines also had the range required, but not the speed, to operate in support of surface warships in the vastness of the Pacific Ocean but were otherwise unremarkable. Furthermore, the design compromises required to carry, deploy, and recover aircraft made these novel submarines decidedly inferior submarines in essentially every other respect – the added weight and volume were not, in other words, without cost or trade-offs. Not only did aircraft-carrying submarines not reflect any fundamental advances in submarine design, propulsion technology, sensing, etc., but the marginal gains in open water scouting capabilities offered should be understood in terms of a context in which large cruisers and battleships also carried a small number of floatplanes for scouting and liaison purposes and in a context in which Japan employed land-based aircraft from Japanese-controlled islands in

[29] David Evans and Mark Peattie, *Kaigun: Strategy, Tactics, and Technology in the Imperial Japanese Navy, 1887-1941* (Annapolis, MD: Naval Institute Press, 1997)

the Pacific as well as increasingly capable fixed-wing aircraft from a progressively more capable and growing fleet of aircraft carriers.

What is notable about the aforementioned efforts toward conceptual innovation is that none of these made any meaningful attempt to address the fundamental limitations of submarine technology in the Interwar Period – the inescapable reality that all but the smallest of submarines, which had very limited range and endurance and which were only suitable for brief operations in coastal waters, were submersibles by virtue of enduring limitations in propulsion technology. Fleet submarines, cruiser submarines, and aircraft-carrying submarines were innovative concepts for how to use much larger long-range submersibles that were developed *despite* enduring technological limitations and none of these developments engendered substantial gains in the fielded military capabilities of adopting states.

8.6 Increased Competition in the Anti-Shipping Role

Although it is important to examine the effects of attempted and realized advances in submarine technology, it is also important not to develop a myopic focus on submarines. After all, militaries did not *have to* pursue submarines in the Interwar Period, and the fact that all major militaries invested in submarines highlights what submarines still offered in terms of capabilities despite the enduring limitations of submarine technology and the mixed results of the employment of submarines in the First World War. To develop a holistic appreciation of the military effects of advances in submarine technology during the Interwar Period, it is productive to consider whether there were newly available substitutes for the roles that submarines could undertake. If such substitutes can be found,

then we have evidence of other technological advances moderating the effects of advances in submarine technology on fielded military capabilities.

While submarine technology largely plateaued throughout the Interwar Period, submarines faced growing competition as a means of torpedo attack against surface warships and merchant ships from an unlikely source – fixed-wing aircraft. As explained in previous chapters, the advent of self-propelled torpedoes in the 1870s made the concept of a submarine more appealing to militaries but militaries did not require submarines to employ self-propelled torpedoes – torpedo boats, destroyers, and larger ships including battleships could all employ torpedoes essentially indistinct from those employed on submarines. Militaries eventually adopted later, more practical submarine designs from the late 1890s onwards, and submarines became an established component of fleets by the start of the First World War, but this was the result of the interplay of a demand-side factor – a requirement for a more survivable submersible torpedo boat – and supply-side factor – the pairing of internal combustion engines with electric batteries and an electric motor in a manner that made a self-propelled submarine more practical (a dynamic cemented by the application of more efficient diesel engines to submarines by circa 1910). Even after submarines became increasingly practical from the early twentieth century onward, militaries continued to equip surface ships of all types with torpedoes and continued to employ large numbers of torpedo-equipped destroyers and torpedo boats, and this dynamic continued throughout the Interwar Period. By the 1930s, militaries began to productively substitute for submarines in certain roles, including the coastal defense role for which submarines were first developed and adopted, with torpedo-armed fixed-wing aircraft, and this moderated the military effects of Interwar

advances in submarine technology in much the same manner as torpedo boats and torpedo-armed destroyers had in the ca. 1870-1910 timeframe.

Fixed-wing aircraft, primarily equipped with a single torpedo, were employed on a very limited scale in anti-shipping roles during the First World War. Given the limited payload, speed, and range of aircraft, as well as the limited means of navigation, communication, and sensing available to the onboard human operator(s), the approach was more effective in principle than it was in practice particularly as warships could maneuver and employ machine guns to deter the more accurate release of torpedoes by aircraft at shorter distances from the target ship. Over the Interwar Period, leading militaries pursued fixed-wing naval aviation in various forms whether land-based aircraft, carrier-based aircraft, or seaplanes of various types (as well as lighter-than-air aircraft – airships – some of which were intended to carry "parasitic" fixed-wing aircraft of their own and serve as airborne aircraft carriers). By the early 1930s, leading militaries deployed growing numbers of dedicated torpedo bombers employing purpose-built aerial torpedoes. Using radios to coordinate attacks with multiple aircraft, militaries could, in principle, undertake large-scale torpedo attacks from multiple axes against both individual surface ships and formations of surface ships and could, of course, do so without employing a single submarine.

Although advances in aircraft technology are rarely discussed in the same breath as advances in submarine technology, it is important to recognize that torpedo-equipped aircraft were superior to submarines in many respects as a means of employing torpedoes against surface ships. Aircraft were, at least in permissive atmospheric conditions (which Interwar Period submarines also required to be effective whilst primarily operating on the

surface), far better positioned to detect enemy ships and quickly transmit the location of enemy ships over radio. If equipped with one or more torpedoes, an aircraft could attack a detected ship itself. Alternatively, it could wait for one or more torpedo-armed aircraft, surface ships, or even a submarine directed to a ship through radio communications, to undertake the attack. The (different, generally smaller, and lighter) unguided torpedoes employed by aircraft were, of course, no more and no less accurate than those employed by submarines and surface ships, but the speed and maneuverability of an aircraft made it a superior platform for targeting a moving ship and re-engaging a target ship if the initial torpedo missed. Torpedo-armed aircraft did, of course, have major limitations. Aircraft were vulnerable to ground- and ship-based air defense artillery as well as gunfire from other aircraft. Torpedo-armed aircraft also lacked the endurance to patrol an area for multiple hours let alone multiple days in the manner of a submarine and aircraft operations were highly indiscreet and thereby attracted adversary aircraft. As such, torpedo-armed aircraft were more effective than submarines in some roles and contexts and less effective in others, and aircraft did not, therefore, fully substitute for submarines.

To be clear, the argument being made here is not that fixed-wing aircraft armed with torpedoes were better than submarines in all roles or that militaries should not have pursued advances in submarine technology in the Interwar Period. Rather, the argument is simply that advances in fixed-wing aircraft technology increased the competition that submarine technology faced in anti-shipping roles, and this affected the magnitude of capability gains that could be engendered by advances in submarine technology. As explained earlier, this was not a new dynamic, but torpedo boats were largely restricted to defensive operations in one's own coastal waters and larger surface warships were too

vulnerable to detection and attack to operate in an adversary's coastal waters not least in the face of torpedo-armed aircraft. Torpedo-armed aircraft could, however, be employed in both offensive and defensive roles in coastal waters and could also operate against surface ships in open waters. As a result, militaries increasingly diversified their capabilities to attack surface ships over the Interwar Period. At the same time, the growing prominence of torpedo bombers had a rarely observed effect on the trajectory of submarine technology – it reduced demand among leading militaries for smaller coastal submarines that were optimized for coastal defense roles. Stated differently, by the mid-1930s, leading militaries primarily built and operated larger, longer-range submarines optimized for longer-duration operations in open waters. This was, of course, a multifaceted dynamic that gradually unfolded over the Interwar Period, but the point stands that fixed-wing aircraft, whether ground- or ship-based, became increasingly important in coastal defense roles as both scouts and combatants and this moderated the effects of advances in submarine technology on the fielded military capabilities of states.

8.7 Developments in Anti-Submarine Countermeasures

Although there were no major breakthroughs in submarine technology during the Interwar Period, the military effects of this dynamic were balanced by the fact that there were also no major breakthroughs in countermeasures to submarines. Gradual advances in submarine technology were as such balanced by gradual advances in countermeasures, technological and otherwise. This measure-countermeasure competition was, like submarine development, dampened by the limited resources available to leading militaries throughout much of the Interwar Period but was also shaped by the interrelated

realities that submarines primarily operated as surfaced submersibles, that submarines primarily were slower than surface warships while surfaced and far slower while submerged, that submarines were vulnerable to gunfire on the surface, and that submarines were vulnerable to depth charges while submerged. In other words, the measure-countermeasure competition in the Interwar Period was very much a continuation of that which transpired during the First World War.

To appreciate the implications of advances in countermeasures over the Interwar Period, it is important to recognize that the bar for effective countermeasures was low, a dynamic vividly demonstrated during the First World War. Not only were submarines exceptionally vulnerable to damage that risked the catastrophic failure of the pressure hull, but militaries did not necessarily need to damage or destroy a submarine in order to neutralize the threat that it posed. Rather, the primary objective was to disrupt the operations of a submarine so as to prevent it from successfully prosecuting targets with a finite supply of unguided torpedoes of limited range. With the oceans being vast and approaches to ports being well monitored and defended (a role for which fixed-wing aircraft were particularly well suited), submarine crews faced major challenges in detecting and tracking potential targets in open water – the open water search problem – let alone damaging or sinking ships. Moreover, the disparities in the maximum speed of submarines relative to surface ships meant that submarines had to operate on the surface to keep up to position themselves for an intercept, thereby making the submarine vulnerable to detection and gunfire, and, when submerged, the submarine had to wait for an opportune ambush scenario to unfold as it could not give pursuit.

Throughout the Interwar Period, naval guns and depth charges remained the primary means through which surface ships could attack submarines. Although naval guns, particularly fire control for naval guns, was an area of technological advancement during the Interwar Period, the bar for damaging an unarmored submarine operating on the surface, let alone forcing the terrified crew to break off an attack and dive, was very low such that even older warships with less accurate guns constituted a potent threat to submarines operating on the surface. Gunfire was not, however, practical against submerged submarines (although a detected submarine operating at periscope depth risked major damage from gunfire), and depth charges remained a core part of anti-submarine capabilities.

During the First World War, depth charges were largely an ad hoc development. Over the Interwar Period, however, depth charges became a distinct type of military equipment and were deployed in increasingly refined forms. The effects of developments in depth charges were nevertheless limited because these were little more than underwater analogs to unguided gravity bombs set to detonate at a preset depth. While a surface ship could carry and employ a large number of depth charges, the effectiveness of the approach was limited by the sheer volume of water that would have to be saturated to damage a moving submerged submarine and detonations sufficiently proximate to damage a submarine were low probability events. In large part, this reflected the fact that Interwar depth charges were essentially dropped overboard from the stern and were not projected laterally or ahead of a surface ship. As a result, a surface ship had to position itself slightly ahead of a submerged submarine for depth charges to be effective. While submerged submarines were much slower than surface ships and notwithstanding the fact

that even a miss could result in a detonation that damaged the submarine hull (i.e., result in leaks and/or mechanical failures), depth charges, being limited in terms of both accuracy (against a moving target) and availability (ships could only carry and deploy a limited number of depth charges), were not very effective in damaging or destroying submarines. Nevertheless, even decidedly imperfect depth charges could serve as an effective deterrent to submarine crews who could not complacently view submerged operations as being equivalent to operating in a sanctuary. Coupled with advances in sonar technology, with which the employment of depth charges interfered, depth charges posed a credible threat to submarine operations throughout the Interwar Period.

While submarines could escape surface ships deploying depth charges, crews had to be ever considerate of the charge levels of their batteries, and here depth charges posed a major if mostly latent threat. If its batteries were close to depletion after diving, a submarine would have no choice but to surface and be subject to naval gunfire. As a result, however ineffective individual depth charges were, these still imposed virtual attrition on submarines as only a reckless crew would prosecute an attack against an armed ship with a depleted battery. A prudent, non-suicidal submarine crew would, in other words, disengage and allow potential (armed) targets to pass out of torpedo range if it had a depleted battery. Battery management therefore became central to the operation of submarines and, in the context of the rapidly increasing prevalence of aircraft equipped with radios capable of reporting the location of a surfaced submarine, atmospheric conditions permitting, the practical effect was that over the Interwar Period, surface operations were increasingly a nighttime endeavor for submarines when operating within

range of land-based aircraft (the range of which also progressively increased over the Interwar Period).

Although there was no major breakthrough in countermeasures to submarines during the Interwar Period, a variety of technological developments created far more challenging operational conditions for submarine crews. Perhaps the single most consequential area of development was, as previously indicated, radio technology in general and radio-equipped fixed-wing aircraft in particular.

By 1914, leading militaries employed radiotelegraphs on their larger warships. By the Interwar Period, radio was increasingly ubiquitous such that by the early 1930s, radios became standard on surface vessels and aircraft of all types. Given the enduring limitations of propulsion technology, submarines were, in effect, submersibles that primarily operated on the surface and could only briefly submerge to approach or escape an enemy ship or otherwise evade detection. A surfaced submarine, whilst smaller than most warships and tiny in the vastness of the open waters of marginal seas let alone the world ocean, was susceptible to visual detection by surface ships, whether merchant or naval, as well as aircraft and coastal observers. As a result, the effectiveness of submarines was sensitive to developments affecting the likelihood that a submarine could be detected even when it was far from any potential targets.

During the First World War, submarine operations were aided by the fact that that until late in the war, most fixed-wing aircraft were not equipped with radios and the same could be said of the most numerous types of ships, whether fishing trawlers or small naval patrol boats, that were, by virtue of their sustained presence at sea, best positioned to detect submarine activity. And even if an aircraft or a naval patrol boat were to report

the location of a moving submarine, radiotelegraphy was far from an ideal means of communication, and it would take time to pass on this time-sensitive information to naval forces that could target a detected submarine. As a result, a moving submarine, whether surfaced or submerged, could displace many kilometers by the time that naval vessels arrived at its last known location. By the 1930s, however, this situation changed as fixed-wing aircraft had the speed to quickly approach an area in which a submarine sighting was reported and, upon confirming the presence of a surfaced submarine, could attack the submarine – which would force the submarine to submerge and utilize scarce battery capacity – and/or report the location to other ships – and other aircraft – in the area. A radio-equipped aircraft could also loiter over the area to track the movement of a surfaced submarine, which was practically indistinct from other surface ships, until warships arrived to hunt for the submarine more persistently. If the submarine dived to evade detection by the aircraft, limited battery capacity meant that it was likely to resurface no more than 50-100 kilometers away from its last known location, and radio-equipped aircraft were well positioned to search the area for a submarine that was, after diving, forced to eventually operate on the surface in order to recharge its batteries.

With surface warships having the endurance required to hunt for submarines but being deficient in terms of sensing and with aircraft being deficient in terms of endurance but being well positioned to detect surfaced submarines, radio-equipped surface warships, and aircraft could cooperate and in so doing could impose considerable virtual attrition on submarines that, given enduring technological constraints, were forced to primarily operate on the surface. For this reason, advances in fixed-wing aircraft technology and advances in radio technology together considerably offset the military effects of advances

in submarine technology over the Interwar Period. Importantly, this was an intrinsic effect of the deployment of radio-equipped aircraft even if a state did not deliberately develop or deploy aircraft for anti-submarine roles. Stated differently, radio-equipped fixed-wing aircraft posed a latent threat to submarines, one that no submarine crew could afford to discount, and aircraft sightings, against which a surfaced submarine had no practical defense, became yet another cause for a submarine to submerge over the Interwar Period and this disrupted submarine operations and thereby imposed virtual attrition.

As with many technological advances, it is important to recognize that advances in radio technology had multifaceted effects on the fielded military capabilities of states. While radio-equipped fixed-wing aircraft posed a growing threat to submarines, advances in radio technology also improved the effectiveness of submarines. As explained earlier, radios allowed submarines to better coordinate their operations with surface ships, land-based headquarters, and, no less importantly, with other submarines. This allowed submarines to transmit and receive the reported locations of enemy shipping and, in terms of intra-submarine communications (whether direct and/or through intermediaries at land-based headquarters), coordinate attacks on groups of surface warships and convoys of merchant shipping. And while radio-equipped fixed-wing aircraft posed a growing threat to submarines, the same aircraft could not only help submarine crews locate and track potential targets but could also quickly surveil coastal waters and report enemy activity to a land-based headquarters that could in turn direct submarines, torpedo boats, surface ships, and armed aircraft to the area.

8.8 Advances in Underwater Sensing Technology

Although active sonar was first developed during the First World War, it would not enter military service until the Interwar Period. Following its introduction, active sonar developed gradually and became more refined but was not characterized by substantial breakthroughs. Even so, by the late 1920s, a growing number of British and American, and later Japanese, warships were equipped with active sonar systems. As a result, the relevant ships in these militaries acquired the nominal capacity to detect and, as such, attack, submerged submarines, something that was not possible during the First World War. Much like active sonar, passive hydrophone technology, which was also developed during the First World War, was not an area of technological development characterized by substantial breakthroughs over the Interwar Period. Nevertheless, hydrophones increasingly became standard equipment on both surface ships and submarines over the 1930s. With hydrophones both facilitating underwater sensing for submarines otherwise reliant on periscopes and offering surface ships a means of detecting submerged submarines, this advance in underwater sensing technology reshaped measure-countermeasure dynamics in undersea warfare.

In principle, these advances in underwater sensing technology offered militaries substantial gains in their anti-submarine capabilities. In practice, however, both active sonar and passive hydrophones had major limitations and the net effects of advances in underwater sensing were limited by the fact that submarines primarily operated as surfaced submarines. Although it constituted a major technological achievement, active sonar was only effective against submerged submarines and was wholly ineffective against submarines operating on the surface. Interwar Period active sonar systems were

also highly directional and limited in terms of range. These were, in other words, analogous to searchlights with a limited field of view like a searchlight mounted under the bow of a surface warship – a ship equipped with active sonar was essentially blind (deaf) in most directions. Compounding matters, active sonar inherently gave away the position of a surface warship, and any ships in its vicinity, to a submarine equipped with passive hydrophones. As a result, Interwar Period active sonar systems were primarily effective when used to defend ships under attack by submerged submarines and were not effective for the purposes of monitoring large volumes of water in search of submerged, let alone surfaced, submarines. The two countries that most eagerly developed and deployed active sonar technology, the United Kingdom, and the United States, operated large numbers of submarines that, like all other submarines, primarily operated on the surface and were therefore well positioned to appreciate the limitations of active sonar technology.

Given the limitations of Interwar Period active sonars, surface ships were heavily reliant on passive hydrophones to detect the presence of a submerged submarine. Passive hydrophones, however, had limitations of their own including low accuracy and poor range, which varied with the speed at which a ship was traveling, the hydrodynamic noise generated by a ship's hull as well as other ships in its vicinity, and variant underwater acoustic conditions. Interwar Period passive hydrophones were not, in other words, remotely comparable to the far more sensitive combination of passive sonar arrays supported by computer-based signal processing capabilities that became ubiquitous

during the Cold War.[30] During the Interwar Period, passive hydrophones were a welcome addition to the sensor suites of surface ships but were nevertheless extremely limited and the visual detection of the periscopes of submerged submarines, of active torpedoes, and the wakes generated by submarines (at periscope depth) and torpedoes was of far greater importance. As a result, there continued to be a relatively level playing field in terms of measure-countermeasure dynamics given that the sensory acuity of surface ships was little better than the sensory acuity of the crews of submarines.

While submarines may have been prematurely rushed into frontline service during the First World War with decidedly mixed results, the two decades or so of technological development during the Interwar Period demonstrate that short of one or more technological breakthroughs, submarine technology was fundamentally limited in terms of the capabilities that it could offer states. Although submarine technology became more refined over the Interwar Period, the latest submarines had the same limitations of designs built over the preceding decades. As a result, the bar to effective countermeasures remained low, and Interwar advances in countermeasures created an operating environment that was increasingly hostile toward submarines in the years leading up to the Second World War. All things considered, the two-decade or so long Interwar Period does not provide evidence of advances in submarine technology engendering sizeable and abrupt shifts in the fielded military capabilities of states.

[30] For a comprehensive account of Cold War dynamics, see Owen R. Cote, "The Third Battle: Innovation in the U.S. Navy's Silent Cold War Struggle with Soviet Submarines," *Newport Paper 16,* (Newport, RI: Naval War College, 2003)

Chapter 9: Submarine Technology ca. 1939-1945

Part Two of this dissertation has so far traced the evolutionary history of submarine technology over a period of around 140 years without finding strong evidence in support of the proposition that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. It is, however, possible that technological advances made during the Interwar Period, limited though these were in substance, had a greater cumulative effect. To determine whether this is the case, we need to extend the longitudinal analysis until the end of the Second World War, which was the second large conflict in which submarine technology was tested in combat conditions.

While the Second World War experience with submarines provides further evidence of the tendency of technological change to not engender sizeable and abrupt shifts in fielded military capabilities, this is overshadowed by the tendency of writings on the implications and effectiveness of submarines in the Second World War timeframe to closely examine one of the two primary submarine campaigns during the war. As with the First World War, the effects and effectiveness of the German submarine force loom large but there was nevertheless no single experience from which to draw generalizable inferences about the military effects of advances in submarine technology. Rather, the Second World War experience is very much a tale of two highly contrasting submarine-centric anti-shipping campaigns respectively undertaken by Germany in the Atlantic Ocean and the United States in the Pacific Ocean that were defined by both highly variant measure-countermeasure competitions and highly variant military-geographical contexts.

In the Atlantic, German submarines enjoyed great successes in targeting merchant shipping for a time (most notably the so-called First Happy Time and the so-called Second Happy Time, two periods in which the German submarine force was unleashed against largely unprepared and slow-to-respond adversaries) but were successfully neutralized by the combination of British and American anti-submarine warfare efforts and German material non-competitiveness that together resulted in the profound defeat of the German submarine force following an intense measure-countermeasure competition.[31] In the Pacific, American submarines enjoyed considerable successes throughout the war in the face of ineffective and inadequate Japanese efforts toward anti-submarine warfare and, no less importantly, on the basis of unmatched American material advantages over Japan. Once we consider other factors, such as geography, it is clear that there is not a single set of inferences to be drawn from the Second World War experience as to the effects of advances in submarine technology on the fielded military capabilities of states. The latest in submarine technology could be very effective (as in the American experience in the Pacific) – could engender substantial gains in fielded military capabilities – or could be greatly blunted if not almost entirely negated by countermeasures (as in the German experience in the Atlantic). Given the profoundly important effects of multiple confounding variables on the two primary Second World War experiences with submarines, there is no reason to think that these contrasting characterizations are intrinsically mutually exclusive, but we nevertheless require a

[31] See Bryan Clark, *The Emerging Era in Anti-Submarine Warfare* (Washington, DC: Center for Strategic and Budgetary Assessments, 2015).; John Stillion and Bryan Clark, *What it Takes to Win: Succeeding in 21st Century Battle Network Competitions* (Washington, DC: Center for Strategic and Budgetary Assessments, 2015)

framework for analysis that allows us to distinguish between the effects of technological change and the effects of other factors.

To account for the aforementioned analytical complications, I will not undertake a detailed analysis of submarine operations in the Second World War but will instead begin with an examination of wartime developments in submarine technology. A benefit of the longitudinal approach to analysis undertaken throughout Part Two of this dissertation is that this chapter can continue where the previous chapter on advances in submarine technology over the Interwar Period left off and the analysis of wartime developments highlights how states adapted to wartime pressures to both field greater numbers of submarines and more effective submarines. I will devote most of this chapter to wartime measure-countermeasure dynamics, which highlight the degree to which the capability gains offered by advances in submarine technology could be blunted if not negated, as well as late wartime development efforts that aimed to produce a substantial breakthrough in submarine technology that would allow (German) submarines to regain the advantage at sea. By examining these important aspects of the Second World War experience with submarine technology, I intend to demonstrate that the dynamics described in previous chapters continued throughout the Second World War such that the history of submarine technology and its effects on the fielded military capabilities of states during the ca. 1860-1945 timeframe, which is approximately an eighty-five-year long experience, was very much a continuous experience.

9.1 Wartime submarine development

The Second World War was characterized by intense efforts toward both developing new military technologies and producing unprecedented quantities of existing and new technologies. The experience with submarine technology was in many respects aberrative in that for the most part, militaries continued to operate and build submarine designs that originated in the Interwar Period until the end of the war. There was, as such, nothing in the world of submarine design and construction that could be compared with the massive production runs of successive aircraft and tank designs, and this is reflective of both the inherent complexity of submarines and the long lead times associated with developing, testing, and manufacturing new submarine designs.

While tanks and aircraft, among other military technologies, were also costly and time-consuming to develop, the interplay of demand-side and supply-side factors made it worthwhile for states to design, test, and deploy, for example, a new version of an existing tank design equipped with a larger gun and/or greater armor protection and/or a more powerful engine, or, alternatively, a new version of an existing aircraft equipped with greater armament and/or a more powerful engine. In the world of submarine technology, there was no comparable pattern of successive technological breakthroughs and technological advances could generally be divided into those that could be readily incorporated into existing submarine designs and those which could not.

Whereas the dimensions of, for example, a tank's hull and its turret placed physical limits on the maximum bore diameter of a new, larger, and more powerful gun that it could be equipped with and thereby created demand for the development of new "clean sheet" tank designs, advances in torpedo technology could be packaged into a

torpedo with dimensions compatible with the torpedo tubes of existing submarines. As a result, states did not require new submarines to employ new torpedoes in the same manner that they required new tanks to operate tanks equipped with a new and much more powerful gun. At the same time, the designers of tanks could replace a turret or attach equipment to the exterior of a tank's hull in a manner that the designers of submarines could not. As a result, some technological advances, generally those that took up considerable weight or volume and which affected the delicate balancing of density in a submarine, required new "clean sheet" designs and considerable research and development efforts followed by testing before these could be manufactured and fielded in large numbers. Given that submarine technology had largely plateaued in the years leading up to the Second World War in the absence of major technological breakthroughs including the absence of new propulsion systems, submarine technology largely continued to plateau in the early 1940s even in the context of a war otherwise characterized by intense and unprecedented levels of research and development and production activity.

Given the above considerations, it is unsurprising that the two countries that built and operated the largest numbers of (large) submarines during the war, which were also the countries that undertook the most notable submarine-centric campaigns, Germany and the United States, both primarily built and employed submarine designs that existed before these countries respectively went to war. Germany continued to produce the late Interwar Type VII and Type IX designs in progressively more capable and refined variants until the end of the war. The United States, for its part, built the closely related Gato, Balao, and Tench classes during the war, and the Gato-class, the first of which was

laid down in October 1940, was the product of nearly two decades of experimentation with various designs for a fleet submarine over the Interwar Period. Other countries, most of which had a less acute demand for new submarines than Germany and which were decidedly inferior to the United States in terms of industrial capacity, did not pursue radically different new submarine designs (with the notable exception of a small number of very large aircraft-carrying submarines built by Japan late in the war). Stated differently, during the Second World War, militaries tended to build more refined versions of late Interwar Period submarine designs and simplified these designs to expedite and lower the cost of production.

While submarine technology gradually advanced throughout most of the Second World War, what may be characterized as component technologies were subject to more significant technological advances and these engendered significant gains in fielded military capabilities. It is important to recognize, however, that these capability gains were not immediate and typically followed early disappointing wartime experiences that facilitated the correction of shortcomings, thereby resulting in effective combat-ready systems.

One of the most notable areas of wartime technological development was in the area of torpedoes. In the late 1930s, Germany discretely undertook research and developmental efforts aimed at producing guided passive acoustic homing torpedoes that homed in on the acoustic signature of a ship's screws. In 1943, Germany introduced such a version of its late Interwar Period electrically powered G7e torpedo. As with many new technologies rushed into frontline service during the exigencies of wartime, initial combat experiences were disappointing, but the design was progressively improved and became

increasingly effective. While Germany had a technological lead through this new guided torpedo design, the practical implications of this lead were marginal in the context of a war characterized by unprecedented levels of research and development activity. Attesting to this dynamic, in 1943, the United States deployed a copy of a captured German G7e electrically powered torpedo in the form of the Mark 18 and, in the same year, the United States separately deployed an acoustic homing anti-submarine torpedo, the Mark 24 FIDO "mine."[32] Given the dearth of German surface ships to target and the effectiveness of its existing unguided torpedoes against Japanese shipping, the United States, which undertook extensive and well-resourced research and development efforts of its own, did not follow Germany in deploying a passive acoustic homing torpedo for use against surface ships.

Although new torpedoes increased the capabilities of Second World War submarines, it is important to recognize that wartime developmental efforts in the area of torpedo technology were primarily directed toward the refinement of existing unguided wet-heater torpedoes. This reflected the unanticipated reality that both the unguided wet-heater German G7a and the unguided wet-heater American Mark 14, the most widely employed torpedoes of these countries' respective submarine forces, were inexplicably unreliable in the opening years of war as a result of inadequate peacetime testing as part of the development, testing, and fielding processes.[33] By 1943, however, late Interwar Period unguided wet-heater torpedoes were perfected and highly reliable armaments, and

[32] Thomas Wildenberg, "A Sub-Hunting Bloodhound," *Naval History Magazine* 31, No. 5 (2017), https://www.usni.org/magazines/naval-history-magazine/2017/october/sub-hunting-bloodhound.
[33] Nickolay Nedelchev, "The Norwegian Operation and the Torpedo Crisis," Uboat.net, April 2, 1999, https://uboat.net/articles/23.html ; Thomas Wildenberg, "A Sub-Hunting Bloodhound," *Naval History Magazine* 31, No. 5 (2017), https://www.usni.org/magazines/naval-history-magazine/2017/october/sub-hunting-bloodhound.

even lavishly funded wartime militaries focused on increasing the production of existing torpedo designs that worked as advertised over new designs that would likely require many months if not several years of development and testing before being ready for combat use.

Armament was not, of course, the only area of component technology in which submarine technology advanced over the Second World War. Radios, sonar, and, in time, radar, among other electronic systems, for example, were areas of fast-paced technological development, and leading submarine operators such as Germany and the United States cycled through multiple radio, sonar, and radar designs in existing and newly built versions of late Interwar Period submarine designs in order to keep apace in intense measure-countermeasure competitions.[34] All things considered, however, the capabilities offered by these submarines were fundamentally limited by their hull designs and propulsion systems such that a German Type IX-class (or a Type VII-class) and an American Gato-class submarine equipped with the latest torpedoes, radios, sonar systems, radars, and so forth were still submersibles that primarily operated on the surface and, given the relatively small number of primarily if not exclusively unguided torpedoes onboard, Second World War submarines were heavily reliant on the employment of deck-mounted guns when and where possible. And while newly built submarines could operate somewhat faster, dive somewhat deeper, and so forth, there were physical limits on what late Interwar submarine designs could offer states in terms of gains in field military capabilities.

[34] Nickolay Nedelchev, "The Norwegian Operation and the Torpedo Crisis," Uboat.net, April 2, 1999, https://uboat.net/articles/23.html; Stillion and Clark, *What it Takes to Win*.

Given the absence of substantial developments in submarine design until very late in the Second World War, which will be examined later in this chapter, it is productive to turn toward measure-countermeasure dynamics to appreciate the pace and magnitude of shifts engendered in fielded military capabilities by wartime advances in submarine technology.

9.2 Wartime measure-countermeasures

Whatever the monadic capability gains offered by technological change, the shifts engendered in the fielded military capabilities of states are necessarily dyadic and are therefore coproduced by the actions taken and forgone by adversaries. It is, therefore, important to closely examine measure-countermeasure dynamics in undersea warfare during the Second World War. In peacetime, as in during the Interwar Period, the effectiveness of countermeasures must be determined through careful analysis. Wartime experiences, however, offer considerable empirical evidence of the effectiveness of countermeasures which, when examined through suitable analytical frameworks, can provide greater insight into the net gains in fielded military capabilities engendered by technological change.

Whereas the First World War was primarily characterized by generally ad hoc countermeasures, both technical and organizational-doctrinal, against submarines, the Second World War experience can be characterized as countermeasures galore. It is, however, important to recognize that no single countermeasure was very effective and that it was the cumulative effects of the combined employment of an increasingly diverse array of countermeasures that allowed the United Kingdom and later the United States to

effectively counter the unprecedentedly large German submarine force. Moreover, as with the First World War, the bar to effective countermeasures was low and militaries did not necessarily need to damage or destroy submarines. Rather, militaries needed to disrupt the operations of submarines and impose virtual, if not actual attrition, and thereby blunt if not negate the effects of deployed submarines and, more generally, advances in submarine technology, on the balance of fielded military capabilities.

9.2.1 Convoys and Escorts

As with the First World War, submarines in the Second World War faced a major open-water search challenge. As a result, while a submarine could in principle, sink a ship, this was easier said than done as it would first have to find the ship and then position itself to successfully attack it with unguided torpedoes. With submarines being relatively slow, limited in terms of sensory acuity, and having a finite number of torpedoes onboard, reducing the probability that a submarine would encounter a ship to attack in the first place offered a straightforward means to undermining the effectiveness of submarines. By operating in convoys, submarines were likely to encounter open waters with a lower density of maritime traffic and even if a submarine detected an ostensibly target-rich convoy, there was strength in numbers as no submarine carried enough (unguided) torpedoes to sink large numbers of moving ships that could invariably advance at maximum speed when under attack. While a submarine could, whilst surfaced, give pursuit to a convoy fleeing the area and attempt to intercept it in order to undertake multiple attacks, merchant ships could be armed, and convoys could be assigned warships to serve as escorts. This would dissuade active pursuit by a trailing surfaced submarine

and force the submarine to rely on its finite supply of torpedoes (i.e., forgo attacks employing deck guns) which, once exhausted, would result in the submarine cutting short its patrol and returning to port.

By reducing the probability of a submarine encountering surface ships and increasing the density of defenses available to merchant ships, escorted convoys dramatically increased the submarine's open-water search problem and increased the likelihood that the subset of submarines that encountered a convoy would face counterattack. As a result, while convoying did not eliminate shipping losses and was not, as such, a perfect solution, it nevertheless greatly blunted the effectiveness of even the latest submarines operated by well-trained crews.[35] As wartime experience and operations research led to more heavily defended convoys, submarines not only faced a greater risk of counterattack but were also more likely to be forced to break off an attack after firing an initial spread of (unguided) torpedoes.

In principle, convoying was a straightforward approach offering many advantages. In practice, there was extensive variation in approaches to convoying and escorting convoys. British and American efforts, which (eventually) constituted the gold standard and the template used in this analysis, were exceptionally well-resourced and far more effective than, for example, Japanese efforts toward convoying. Although it is tempting to explain this variation in terms of measure-countermeasure dynamics in undersea warfare, industrial capacity, and resource availability, among other factors, played a central role not only in terms of compensating for wartime losses but also in

[35] Notably, shipping lost to German submarines fell before German submarine losses increased, a dynamic that is reflective of waning submarine effectiveness short of highly effective anti-submarine capabilities that resulted in sunken submarines. See, Clark, *The Emerging Era in Anti-Submarine* Warfare, 4.

terms of increasing the capacity of the merchant fleet and providing large numbers of warships available to serve as escorts. Convoying was not, in other words, a panacea, and the wartime effectiveness, or lack thereof, of convoying was a product of a diverse array of factors unrelated to submarine technology and anti-submarine warfare capabilities and technological advances in these areas. Nevertheless, a submarine's open water search problem and the virtual attrition imposed on submarine forces by anti-submarine warfare efforts including convoying played a central role in determining the effects of advances in submarine technology during the Second World War timeframe. Being slow, vulnerable, limited in terms of endurance, and reliant on the successful employment of a small number of unguided torpedoes, submarine operations were highly susceptible to adversary action.

As warship production increased and as new dedicated ocean-going escorts were developed and built in large numbers, the British and later the American militaries went beyond merely escorting convoys to more aggressively hunting submarines. There were two general and increasingly overlapping approaches pursued by the United Kingdom and the United States. The first (non-chronological) approach pursued by both the United Kingdom and the United States entailed the formation of so-called hunter-killer groups of surface warships to hunt German submarines. While warships could escort convoys, this was an entirely reactive role that forced surface warships to operate at a much slower speed to maintain the convoy formation alongside slower merchant ships. By operating independently of merchant ships, hunter-killer groups could move quickly at sea in response to information on submarine activity determined by a range of sources,

including signals intelligence intercepts of German submarines coordinating over radio to intercept a convoy as a so-called wolfpack.

The second approach involved the use of ship-borne aircraft to support convoy operations. Faced with a shortage of aircraft carriers, the United Kingdom initially adapted its merchant ships to carry aircraft. CAMs, or catapult aircraft merchant ships, launched single-use catapult-launched fighter aircraft not to attack German submarines but the small number of long-range fixed-wing maritime patrol aircraft that Germany used to help its submarine force overcome the immense open water search problem encountered in the vastness of the northern Atlantic Ocean. With the United Kingdom having no aircraft carriers to spare in the opening years of the war, CAM ships offered an important but nevertheless limited and indirect gain in anti-submarine capabilities. The United Kingdom also converted merchant ships into ad hoc aircraft carriers equipped with flight decks. Unlike CAM ships, the aircraft deployed on merchant aircraft carriers, or MAC ships, were reusable and could undertake multiple sorties. This allowed convoys including MAC ships to undertake aerial patrols with radio-equipped aircraft. As dedicated light aircraft carriers and escort carriers became available in increasing numbers later in the war, convoys were protected by larger numbers of armed radio-equipped aircraft that could not only patrol the area in which a convoy was operating but also attack detected submarines. As a result, aircraft-equipped convoys were able to, in effect, increasingly turn the area of open water in their vicinity into as hostile an area operation for surfaced submarines as coastal waters.

Although convoys are often characterized as groups of warships and merchant ships, land-based aviation also played an increasingly important role in protecting

convoys as the war progressed. Large long-range aircraft were limited in number in the opening years of the war and most such aircraft were bombers and were not employed in maritime patrol roles. As a result, convoys, including escorted convoys, did not benefit from the support that could have been provided by radio-equipped aircraft outside of coastal waters. In time, however, dedicated seaplanes and land-based aircraft with greater range were developed and built and these were used to surveil vast swathes of the Atlantic Ocean for submarines operating on the surface and more generally monitor areas that a convoy was expected to traverse.

Fixed-wing aircraft were, of course, limited in terms of endurance but from ca. 1943 onward, the so-called mid-Atlantic Gap, the "middle" section of the North Atlantic previously beyond the reach of land-based coastal patrol aircraft based on either side, was "closed" through the employment of large numbers of patrol aircraft with suitable range. As a result, German submarines undertaking highspeed transits on the surface to intercept convoys based on information transmitted over radio were increasingly vulnerable to detection by radio-equipped aircraft that could, in turn, report the submarine sighting to convoys in the area and to land-based headquarters who could reroute convoys as required. As the aircraft employed in maritime patrol roles improved in terms of range, endurance, payload, and, no less importantly, sensors (i.e., radar), armed patrol aircraft posed an ever-growing threat to submarines that were forced to primarily operate on the surface because of enduring limitations in submarine propulsion technology. As explained earlier, a submarine could, of course, dive and evade attack and further detection but it was severely limited in terms of submerged speed and range and would have to surface within 24 to 48 hours to recharge its batteries. Ultimately, submarines

were not built and operated merely to survive at sea, and the disruption of submarine operations and the formation of so-called wolfpacks by radio-equipped and increasingly well-armed patrol aircraft greatly blunted the effectiveness of the German submarine force.

9.2.2 Active Sonar and Passive Hydrophones

Convoying compounded the search problem faced by submarines and the use of escorts increased the threat faced by a detected submarine. As a result, submarines were less likely to find ships to sink, less likely to be able to attack ships whilst surfaced using deck guns (for which there was more ammunition available onboard than torpedo reloads), and were more likely to attempt to escape the area after launching torpedoes. Together, convoys and escorts undermined the effectiveness of submarines even if the submarines were undetected until the convoy was attacked. Submarines could, however, nevertheless attack with torpedoes while submerged, and convoys and their escorts had no defense until the presence of a submerged submarine was detected, which was often after a torpedo detonated upon impacting a ship in the convoy. Successive advances in sonar technology, however, allowed convoys to actively search for submerged submarines and actively fight back. While submerged submarines often had an upper hand in such encounters, a submarine escaping a convoy's escorts after a successful attack was a submarine that was not sinking any additional surface ships. As a result, the effectiveness of submarines against convoys was highly sensitive to developments in active sonar and passive hydrophone technology.

Active sonar was primarily developed by the United Kingdom and the United States over the Interwar Period, and the former was heavily reliant on active sonar, or ASDIC as it was called in the United Kingdom, to protect the convoys that were formed following the start of the war. In principle, active sonar was a formidable anti-submarine asset but, as explained in the previous chapter, it was useless when submarines were launching torpedoes and firing deck guns at night whilst surfaced, and this was the standard practice of German submarines operating in so-called wolfpacks. As a result, active sonar was of limited effectiveness until later in the war, when other developments, above all the introduction of surface search radars to surface ships and aircraft, made surfaced operations increasingly dangerous for submarines that were, in turn, forced to attack convoys while submerged. Nevertheless, the very existence of active sonar meant that submarines did not have an absolute sanctuary while submerged particularly as multiple sonar-equipped warships could effectively localize a submerged submarine's location and target it with depth charges. While depth charges were inaccurate and largely ineffective, active sonar allowed warships to interfere with a submarine's operations by forcing it to leave the area and, in so doing, draw upon scarce battery charge.

While active sonar receives considerable attention in writings on antisubmarine warfare in the Second World War, it is important to recognize that passive hydrophones also played an increasingly important role in countering submarines. Hydrophones offered submarines a means of detecting surface ships and surface ships a means of detecting submarines in low visibility conditions (sea conditions permitting). While submarines benefitted from the employment of passive hydrophones, the diffusion of

hydrophones on growing numbers of surface warships meant that submarines could no longer find sanctuary by virtue of operating at night, in low visibility conditions, or by remaining submerged. Hydrodynamic noise, machine/engine noise, and the rotation of propellers generated acoustic signatures that alerted surface ships and submarines alike of one another's presence, thereby greatly reducing the likelihood that a submarine could approach surface ships whilst undetected and thereby employ unguided torpedoes to maximum effect at short range.

As sonar technology progressed over the Second World War, it facilitated the development of a new array of armaments that could be employed both by and against submarines. Germany led the way in the development of passive acoustic homing torpedoes, namely the electrically powered G7e, to attack convoys defended with depth-charge projectors. This in turn led to the development and use of acoustic decoys on surface ships such as the British Foxer that attempted to lure acoustic homing torpedoes toward false targets. The United States, for its part, developed the Mark 24 FIDO air-dropped passive acoustic homing torpedo (ostensibly a "mine") to attack submerged submarines. As a result, submarines and surface ships were engaged in an intense measure-countermeasure competition that moderated the effects of wartime advances in submarine technology.

9.2.3 Depth Charge Projectors

Active sonar and passive hydrophones allowed warships to detect submarines but offered no means of damaging let alone destroying a submerged submarine. While surface ships could use searchlights and, in time, radar, to detect a surfaced submarine and direct

gunfire toward its location, thereby forcing it to submerge or risk catastrophic damage to its pressure hull, surface ships were otherwise reliant on depth charges to attack submerged submarines. Depth charges were, by their very nature, inaccurate weapons that primarily existed to present a latent threat to a submarine's crew and were therefore far more effective in terms of imposing virtual attrition than actual attrition. With active and passive sonar offering surface ships the ability to detect submerged submarines, militaries pursued what amounted to depth charge projectors to supplement ordinary depth charges. This became a particularly important undertaking because active sonar was a forward-looking sensor with a limited field of view (early wartime active sonars were also unable to determine target depth). Depth charges, however, had to be dropped over the stern such that while active sonar allowed a warship to pursue a submarine, the warship would have to undertake what amounted to "blind bombing" with active sonar once it lost contact with the moving and possibly maneuvering submerged submarine that was supposed to be underneath it. To address this issue, militaries required what amounted to a depth charge projector. The United Kingdom developed the Hedgehog, a depth charge projector that was essentially a spigot mortar with a range of around 200 meters, and later developed a separate system, the Squid.[36] The United States, for its part, developed a rocket-powered system, the Mousetrap. When used alongside active sonar, these depth charge projectors posed a new and very different threat to submarines.

[36] "Hedgehog Weapons," National Museum of the U.S. Navy, No Date, https://www.history.navy.mil/content/history/museums/nmusn/explore/photography/wwii/wwii-atlantic/battle-of-the-atlantic/anti-submarine-warfare/hedgehog.html ; "Mousetrap Weapons," National Museum of the U.S. Navy, No Date, https://www.history.navy.mil/content/history/museums/nmusn/explore/photography/wwii/wwii-atlantic/battle-of-the-atlantic/anti-submarine-warfare/mousetrap.html

Prior to the introduction of depth charge projectors, an undetected submarine could approach a ship or a group of ships, launch a spread of torpedoes, and, if so inclined, reload to launch a second spread. With its presence unavoidably heralded by the detonation of the first spread of torpedoes (assuming all the torpedoes did not miss), the submarine was inherently in greater danger, but it could submerge/could remain submerged provided that it had sufficient battery charge. As a result, even if warships approached its location to drop depth charges from the stern of the ships, the submarine could maneuver while submerged at its maximum speed to escape the area. Stated differently, a submerged submarine with the advantage of surprise was well positioned to launch a spread of torpedoes and escape without warships having a realistic opportunity to deploy depth charges. With the advent of depth charge projectors, however, escorts could immediately deploy depth charges in the area of the submarine's location as determined by active sonar and/or passive hydrophones. The submerged submarine therefore had no time to lose in escaping the area and faced the additional threat that faster surface ships could give pursuit and use depth charge projectors to attack the submarine within the field of view of active sonar. With depth charge projectors like Hedgehog employing contact fuses, the combination of advances in sonar technology and depth charge projectors constituted a qualitatively different threat to submerged submarines that threatened to quickly turn the would-be hunter into the hunted. As a result, not only did depth charge projectors amplify the effects of advances in sonar and impose actual attrition on submarines, but this advance in anti-submarine warfare capabilities also imposed virtual attrition by inducing submarine crews to become even more risk averse.

9.2.4 Signals Intelligence

As explained in a previous chapter, signals intelligence played an important role in anti-submarine warfare efforts and military operations more generally in the First World War. For militaries, radio technology offered considerable capability gains by allowing units, including individual submarines, to coordinate, but communication also created new vulnerabilities by giving away the location of the transmitter and, depending on the capacity of one state to encrypt information and the capacity of another state to decrypt said information, the content of their communications. All things equal, the greater the distances involved between units attempting to communicate with one another, the greater the reliance upon radio communication. While militaries could use the medium of copper wire to communicate more securely over land, communications over water necessitated the employment of radio. As a result, militaries turned once again to signals intelligence efforts to both find surface ships to attack and submarines to defend against.

Given the open water search problem, submarines were highly reliant on radio communications to receive information on enemy shipping activity from other submarines, surface ships, aircraft, and, no less importantly, land-based headquarters with access to signals intercept and decryption capabilities. At the same time, surface warships and convoys of merchant ships were similarly reliant on radio communications to coordinate operations – including forming convoys – and transmit and receive information on enemy submarine activity. With military communications universally encrypted by the start of the Second World War, undersea warfare was increasingly shaped by an intense measure-countermeasure competition between communications

intercept and decryption technology (and practices) on the one hand and communications security and encryption technology (and practices) on the other.

The Second World War witnessed a wide range of developments in signals intelligence capabilities.[37] Surface ships and aircraft received radio direction-finding equipment and were therefore able to add signals intelligence to the growing array of sensors that could be deployed at sea to detect submarines. Whereas submarine hulls and onboard machinery were not areas of substantial technological breakthroughs during the Second World War, the apparatus used to encrypt, decrypt, transmit, and receive radio signals at the start of the war were thoroughly obsolete by the end of the war and were progressively replaced by more advanced systems. Toward the end of the war, submarines were beginning to receive burst radio communications systems that were more difficult to intercept, Germany went as far as to deploy a Very Low Frequency (VLF) long-wavelength radio communications systems that could transmit messages to submerged submarines (submarines still had to surface to transmit radio emissions and could not do so through VLF outbound transmissions), and signals intelligence capabilities were advancing on the basis of advances in computer technology, among other things. Signals intelligence was, in other words, itself an area of intense measure-countermeasure competition even as it took on an increasingly central role in the measure-countermeasure competition in undersea warfare.

Although British and American signals intelligence efforts during the Second World War receive most of the attention by virtue of being the most successful, it is important to recognize that all leading militaries had their share of successes and failures

[37] See, for example, David Syrett, *The Battle of the Atlantic and Signals Intelligence: U–Boat Situations and Trends, 1941–1945* (London, UK: Routledge, 2002)

and that breakthroughs in signals intelligence, such as the decryption of messages transmitted through the German Enigma machine, were far from guaranteed and lay at the intersection of contingency and tremendous ingenuity and effort.[38] As a result, it is productive to view advances in signals intelligence capabilities as inherently making all submarines communicating over radio more vulnerable to detection (i.e., irrespective of cross-national variation in signals intelligence capabilities). If faced with an adversary with formidable anti-submarine capabilities and increasingly widespread radio direction-finding capabilities, which can be effective independent of the capacity to decipher intercepted encrypted messages, then coordinated submarine operations, such as Germany's so-called "wolfpacks," were at great risk, and submarines were, therefore, more productively employed in the less effective role of a solitary hunter that unavoidably encountered a profoundly challenging open water search problem. If, however, submarines operated against an adversary with weaker anti-submarine capabilities, then a submarine force, such as that of the United States operating in the Pacific, did not have to extensively coordinate over radio to the same extent as the German navy and the effects of advances in signals intelligence were, therefore, more limited. As with all countermeasures against submarines, signals intelligence capabilities were not independently decisive.

9.2.5 Radar

Although active sonar, passive hydrophones, signals intelligence capabilities, and fixed-wing aircraft all offered militaries means of discerning the location of submarines, none

[38] Jak P. Mallmann-Showell, *German Naval Codebreakers* (Annapolis, MD: Naval Institute Press, 2003)

provided a sensor that was capable of surveilling large areas and tracking surfaced submarines in real-time. Radar, however, provided such a capability and, over a period of several years, entered what became the final nail in the coffin of the universal decades-old concept of a submarine as a submersible that primarily operated on the surface.

Radar was independently developed in multiple states over the 1930s but did not become operational on a large scale until ca. 1940.[39] Early radars had many limitations, and all Interwar Period radars were low frequency/long wavelength designs that were particularly unsuitable for the wide range of roles to which later radars would be applied including the detection of small surface vessels such as a surfaced submarine. By the late 1930s, multiple countries were independently experimenting with large radar designs for use on land in air defense roles and for use on the largest surface warships (i.e., battleships and heavy cruisers) to detect aircraft and surface ships. Early radar designs were very large and cumbersome but by ca. 1940-1941, a range of more advanced and more compact higher frequency/shorter wavelength radar designs were introduced including both the first air-to-air, or airborne intercept, radars that could be mounted on aircraft as well as the first more compact naval surface search radars that could be employed on smaller ships. Importantly, in 1940, the United Kingdom made a technological breakthrough with the development of the cavity magnetron, a device that enabled the development of far more compact and higher-resolution centimetric radars that operated at higher frequencies/shorter wavelengths. Although centimetric radars would not see frontline use until early 1943, this technological development dramatically

[39] Raymond Watson Jr, *Radar Origins Worldwide: History of its Evolution in 13 Nations Through World War II (*Bloomington, IN: Trafford Publishing, 2009); Gaspare Galati, *100 Years of Radar* (New York, NY: Springer, 2015); S.S. Swords, *Technical History of the Beginnings of Radar* (Stevanage, UK: Institution of Engineering and Technology, 2008)

affected the measure-countermeasure competition in undersea warfare in ways that few could have anticipated in the Interwar Period.

From ca. 1939 to 1942, submarines could operate in relative safety on the surface in open water and could do so with near total impunity at night. While surfaced submarines could be detected through the use of hydrophones and signals intelligence efforts, hydrophones were exceedingly limited in terms of range, and signals intelligence not only took time to develop but submarines did not continuously transmit radio messages. And because submarines were highly – essentially constantly – mobile, unlike ground combat units that made use of a radio, the last known location of a submarine only indicated the approximate area in which a submarine was operating and not its location in real-time. While intercepted submarine communication could indicate that a convoy was detected and would face attack in a given area, the ships in the convoy had to know the precise location of a submarine to know when to undertake evasive maneuvers and when to attempt an attack on an approaching submarine. Aircraft patrols could help compensate for the limitations of hydrophones and signals intelligence capabilities, but the only sensors available to aircraft prior to the introduction of suitable radars were human eyeballs and aircraft could not therefore detect submarines at night or in low visibility atmospheric conditions. Surface ships were similarly heavily reliant upon human eyeballs assisted by optical instruments to detect submarines and, if encountering a submerged submarine, would have to be on watch for a protruding periscope that would give warning of a potentially impending torpedo attack. As a result of these limitations in sensors that could be used to detect submarines, the concept of the submarine as a submersible that primarily operated on the surface remained viable.

With the advent of suitable radars in the early 1940s, surface ships acquired a sensor capable of detecting surfaced submarines over several kilometers that was equally effective both day and night. When more compact and higher resolution radars became available for use on aircraft, long-range fixed-wing patrol aircraft could similarly detect surfaced submarines at distances of several kilometers (some aircraft also received powerful searchlights to help determine the precise location of a surfaced submarine so that the detected submarine could be attacked at night with unguided weapons).

With most unguided torpedoes having a maximum range of little more than ten kilometers – at lower speed settings – and with a little more than two dozen torpedoes available on even the largest submarines; submarine commanders had every incentive to approach a target ship as close as possible to maximize the probability of kill against moving targets that were capable of maneuvering. To attack a convoy composed of multiple moving ships, a submarine commander would preferably attack on the surface at night. By so doing, a surfaced submarine could give pursuit to a convoy and launch multiple torpedoes. With the advent of ship-based search radars, however, submarines could no longer realistically approach a surface ship or a convoy whilst surfaced even during the night given the threat posed by gunfire. As a result, a submarine had to approach its intended targets at a slower speed while submerged and this posed two major challenges that undermined the effectiveness of submarines. Before the advent of ship-based radar, surfaced submarines operating in suitable lunar conditions could pursue a convoy whilst operating on the surface during the night and therefore undertake multiple attacks. Given the poor energy density and specific energy of electric batteries and the fact that battery charge was a scarce commodity desperately required to allow a

submarine to escape convoy escorts, a submerged submarine, being far slower than the convoy, could not simply give pursuit. As a result, submarine commanders had fewer opportunities to employ torpedoes before the convoy was out of range and this incentivized launching torpedoes from higher risk positions relative to the convoy that increased the likelihood of being detected by sonar and being attacked with depth charge projectors.

Compounding matters for submarine crews, the combination of long-range patrol aircraft and higher resolution radars allowed patrol aircraft to better detect surfaced submarines even in low visibility conditions and, more importantly, allowed aircraft to detect surfaced submarines at night. Until radar warning receivers could be developed and deployed, radar-equipped aircraft could also directly approach a detected surfaced submarine and attack it with bombs, torpedoes, and gunfire. As a result, submarines, which had to operate on the surface to recharge batteries, operate at maximum speed, and reposition to intercept convoys based on information transmitted over radio, had to be on constant guard for enemy aircraft and had to quickly dive when aircraft were sighted. This severely interfered with submarine operations and inflicted virtual attrition even if submarines evaded aerial attack without incurring damage.

While coastal waters, being areas characterized by navigational hazards, were never safe areas for larger ocean-going submarines, radar-equipped aircraft posed a particular threat to submarines departing or approaching a port irrespective of the time of day. Without friendly fighter and interceptor aircraft to patrol the airspace in areas adjacent to the ports from which submarines operated, submarines heading out to sea or returning from a patrol could no longer transit at maximum speed whilst surfaced and

instead had to undertake slower, often more circuitous routes and repeatedly submerge to evade aerial attack. The aerial application of the latest advances in radar technology therefore had more far-reaching effects than the employment of radar on surface ships and interfered with submarine operations even in areas in which no convoys or surface warships were present.

As submarine sinkings increased, militaries operating submarines against adversaries employing radar-equipped ships and aircraft pursued countermeasures. Although all radars are intrinsically susceptible to jamming activity, this was not a productive approach for a submarine as the priority was to dive to evade attack from a warship or a plane and jamming would require the submarine to remain on the surface where it would be exposed to gunfire without any concealment from darkness. Moreover, even if a submarine were to successfully jam a radar, the emissions of the jammer would give away the location of the submarine. Militaries therefore pursued two different types of countermeasures directed against the employment of radar.

The first approach was to develop radar warning receivers that would alert the crew of detected active radar emissions. Given the nature of the propagation of electromagnetic waves emitted by radar, the intended target can inherently detect radar emissions at a greater range than the ship or aircraft on which the emitting radar is mounted. When equipped with such systems, a surfaced submarine would receive early warning of the presence of a ship or, more likely, an aircraft equipped with a radar and could prepare to quickly dive. Leaving aside the fact that a radar warning receiver did not improve the ability of a submarine to attack surface ships, this countermeasure had a limited shelf life. Radar technology was experiencing a – never to be repeated – phase of

rapid growth spurt toward maturity in the 1940s such that a radar warning receiver would not work against a new radar operating at a higher frequency/shorter wavelength, thereby necessitating the development of new radar warning receivers. Until such a new radar warning receiver could be developed, tested, built, and fielded, submarines were at a disadvantage, thereby establishing a dynamic in which one side or another could have the upper hand in terms of employing/detecting radar for a period generally no longer than several weeks to several months. This was, in other words, an area of intense measure-countermeasure competition strongly shaped by the capacity of a state to discern new vulnerabilities and develop, test, manufacture, and field updated electronic devices, both radar and radio, at scale.

The second approach to mitigating the effects of radar on submarine operations was to develop surface and air search radars for use by submarines. Although the use of these radars would give away the location of a submarine, it nevertheless reduced the risk of a submarine being caught by surprise, an important consideration given that submarines had to operate on the surface to recharge their batteries. When operating radars, submarines avoided the possibility that their radar warning receivers would not work and also gained greater awareness of ships and aircraft operating in their vicinity.

9.2.6 Countermeasures Summary

By the end of the Second World War, measure-countermeasure dynamics in undersea warfare had reached a point that submarine operators from the late 1930s, let alone the First World War, without the wartime experience would likely have been overwhelmed. Until the early 1940s, submarines could readily operate, primarily on the surface, in much

the same manner as submarines in the 1910s. By 1943, however, multiple factors, only some of which are related to technological change, converged to create a situation in which advances in countermeasures greatly blunted the effects of advances in submarine technology and, absent a breakthrough in propulsion technology, the effectiveness of submarines against responsive, adept, and well-resourced adversaries was greatly diminished.

Individually, the various countermeasures to submarines discussed in this chapter were of non-zero but nevertheless limited effectiveness. Convoys compounded the open water search problem and escorts made it more difficult to attack convoys whilst surfaced, but submarines could, and did, successfully sink merchant ships in convoys even if the majority escaped unscathed. Similarly, sonar and depth charge protectors increased the threat posed by surface ships to submerged submarines, but submarines could, and did, continue to attack both individual ships and ships that were part of convoys. The combination of signals intelligence, the large-scale use of radio-equipped armed aircraft, and, most of all, radar (primarily on aircraft), however, had a far greater cumulative impact on the effectiveness of submarines by depriving submarines of sanctuary whilst surfaced in open water even when a submarine was many kilometers from the nearest surface ship or convoy.

To fully appreciate the scope of measure-countermeasure competition by the end of the Second World War and the effects it had on the military effects of advances in submarine technology, we need to turn to advances in submarine technology made in the final years of the Second World War. While few of the latest and most advanced specimens of submarine technology entered service by the end of the war, these

nevertheless heralded the emergence of what may be productively characterized as "true submarines" – as opposed to submersibles – after decades of sustained if gradual technological development and highlighted just how contingent the observable outcomes of the Second World War were in the area of undersea warfare.

9.9 The Advent of "True" Submarines

Until 1944, essentially every self-propelled submarine ever built intended for roles beyond harbor defense[40] was designed to primarily operate on the surface and only briefly submerge in order to attack enemy surface ships, escape attack from enemy surface ships, or otherwise evade detection. Attesting to this dynamic, until 1944, essentially every submarine ever built had a maximum surface speed that was considerably higher than its maximum submerged speed, which is to say that essentially every submarine built had a hull design that was deliberately optimized for operation on the sea surface as a submersible at the cost of a slower submerged maximum speed. By 1943, however, the dynamics of the multi-year-long measure-countermeasure competition between German submarines and British and American anti-submarine warfare capabilities, above all the combination of signals intelligence, the large-scale use of radio-equipped aircraft, and radar, reached a point in which the large German submarine fleet was for all intents and purposes neutralized in the North Atlantic. With neither the aircraft nor the surface ships required to contest for access to the Atlantic or substitute for its submarine fleet and thereby facing outright defeat at sea, Germany

[40] A very small number of submarines, mostly built by France in the late nineteenth century, were solely equipped with electric motors powered by electric batteries and were, as such, primarily subsurface vessels. These were, however, extremely small and had negligible endurance and were as such restricted to the role of harbor defense.

began to modify existing submarines and develop entirely new submarine designs that were intended to primarily operate whilst submerged. This marked a turning point in submarine technology, one that laid the foundations for the advent of "true" submarines as opposed to torpedo- and gun-armed submersible surface combatants.

9.9.1 Understanding the Context in Which Snorkels Were Developed

To understand the advent of "true" submarines and the implications of this technological advance on fielded military capabilities and the Second World War measure-countermeasure competition in undersea warfare, we must first examine the deceptively simple appearing technological advance that was the adoption of snorkels for submarines. By so doing, we will be better positioned to understand the context in which two radically new snorkel-equipped German submarine designs, the Type XXI and Type XXIII, were developed and deployed, and the implications that these new designs did and did not have on the late war naval balance.

As explained in previous chapters, by 1900, designers in multiple countries converged on a dual propulsion system for submarines that utilized an electric motor powered by electric batteries for submerged propulsion and an internal combustion engine of some type – later exclusively diesel engines – as both a prime mover for operations on the sea surface and an electrical generator to recharge the onboard electric batteries that would power the submarine whilst it was submerged. The poor energy density and specific energy of batteries, among other technical limitations associated with battery technology throughout most of the twentieth century, meant that submarine designers either had to accept brief submerged endurance or pursue far larger and more expensive submarines that could house a greater number of battery cells. Whilst the latter

approach was appealing, it came at a steep cost and inherently diminishing returns, and such a submarine would still need to operate on the surface to operate the internal combustion engine and thereby recharge the batteries. All things equal, including the power rating of the onboard combustion-powered electrical generator, the greater the number of battery cells of a given battery chemistry onboard a submarine, the longer that a submarine would have to operate on the surface before fully recharging its batteries. Moreover, the larger and heavier the submarine, all things equal, the greater the energy expenditure, whether sourced from a combustion process or electrical batteries, required to propel the submarine to a given speed whether on the surface or whilst submerged. Absent the availability of a radically different propulsion system or a fundamental breakthrough in battery technology, the only practical way to escape these intertwined technological constraints was to develop a means of running internal combustion engines within an essentially submerged submarine (i.e., a submarine that was mostly beneath the water's surface) without asphyxiating the crew. This required (A) a means of drawing in an oxidizing agent, oxygen, from the atmosphere,[41] and (B) a means of ejecting noxious combustion products from the atmosphere contained within a submarine's hull.

As with many military technologies, the conceptualization of a snorkel-like device for submarines predates the development, testing, and use of such a device. Snorkels were most fully developed before the Second World War in the Netherlands, which trialed snorkels on its O-21-class submarines in the late 1930s.[42] Both the United Kingdom and Germany accessed several of these snorkel-equipped submarines in 1940

[41] Alternatively, designers could utilize liquid oxygen or hydrogen peroxide, both of which could be stored within a submarine.
[42] Mark C. Jones, "Give Credit Where Credit Is Due: The Dutch Role in the Development and Deployment of the Submarine Schnorkel," *The Journal of Military History* 69, No. 4 (2005): 987-1012.

but, after testing, neither country adopted snorkels for their existing or new-build diesel-powered submarines even though snorkels self-evidently offered a means through which submarines could primarily operate whilst mostly submerged. At first glance, this dynamic may appear very strange if not nonsensical, but it is important to recognize that several factors disincentivized the adoption of snorkels by Germany, the United Kingdom, and other submarine operators.

First, although the concept of snorkel was relatively straightforward, it was difficult to implement, not least in a wartime context, with the level of difficulty varying with the operating environment for which a submarine was intended. The Netherlands developed snorkels for new-build submarines (i.e., with snorkels integrated into the hull) intended for operations in the tropical waters around the Dutch East Indies. In that environment, unlike the North Sea adjacent to the Netherlands (as well as Germany and the United Kingdom), heat and humidity made surface operations on a submarine extremely uncomfortable for the crew, a dynamic exacerbated by the heat generated by the onboard diesel engine and the risk posed by the effects of condensation on the onboard batteries. While the Netherlands also pioneered the use of air conditioners in its submarines, a submarine could, by means of a snorkel, run its diesel engines while operating just beneath the water's surface without asphyxiating its crew and poisoning the atmosphere contained within the hull.

While submarine operations in the North Sea and North Atlantic were not characterized by heat or humidity, in these environments, unlike the tropical waters around the Dutch East Indies, the regularity of unstable surface navigation, a product of rough seas, meant that militaries had to account for the risk that water would enter the

snorkel while the diesel engines were running. While more refined snorkels included shut-off valves designed to prevent water from entering the engines, this would result in the creation of a partial vacuum within the submarine and result in severe pain in the ears of the crew. In the generally calmer tropical waters around the Dutch East Indies, this was a lesser concern. As with other cases, the military-geographical context in which a technological advance was applied shaped its practicality and, as such, its effects on fielded capabilities.

Environmental factors were not the only barrier to the adoption of snorkels for primarily surface-running diesel-powered submarines in the North Atlantic military-geographical context. Although a snorkel-equipped submarine could operate just below the sea surface, the use of snorkels created new hazards for submarine crews and imposed new restrictions on the operation of submarines. Whereas a typical large ocean-going submarine in service ca. 1940 could sustain a speed of fifteen to twenty knots whilst surfaced and five to nine knots whilst submerged, the use of a snorkel on a converted submarine limited the maximum speed to no more than six knots. As a result, a snorkel-equipped submarine would be less susceptible to detection and, as such, to attack, whilst transiting in open water but would also be far slower. The onboard provisions were, however, unaffected by the existence of a snorkel and the net effect was that otherwise unmodified snorkel-equipped diesel-powered submarines would be limited to shorter duration missions for want of supplies rather than for want of fuel. As a result, otherwise unmodified snorkel-equipped submarines had less time on station to patrol in search of enemy ships and were incapable of pursuing faster enemy ships whilst at snorkel depth to bring said ships within range of their unguided torpedoes (i.e., the submarine still had to

operate on the surface using diesel engines to intercept faster-moving targets). Beyond matters concerning the restrictions on speed imposed by the use of a snorkel, snorkels also restricted the sensory acuity of the crew of an otherwise unmodified diesel-powered submarines as it interfered with the operation of periscopes – as a result of vibrations – and hydrophones. Moreover, the use of a snorkel left a distinct visual signature – a wake – on the sea's surface such that snorkeling did not eliminate the risk of being detected and/or targeted by enemy aircraft.

While a new submarine design centered on the employment of a snorkel could address some, but not all, of these issues, a clean sheet design for something as complex to design and build as an entirely new submarine would likely require several years to develop, test, refine, and finally manufacture in large numbers. Germany did not have such time in the Second World War, and it is also worth stressing that the Netherlands, which pioneered snorkel technology, did not pursue snorkels as a result of concerns about the survivability of its submarine whilst operating on the surface. Throughout the Interwar Period, this was not a major concern for navies, and, in the absence of radar, submarines could always find relative sanctuary on the surface in open water and, could, in a worst-case scenario, shift to nighttime surface operations. As a result, while the Netherlands developed the first snorkel-equipped submarines, it did not develop submarines that were optimized for submerged operations. The hulls of the O 21-class submarines were, like all other large submarines built since ca. 1900, optimized for surface speed and were, as such, considerably faster on the surface when running on the diesel engine than when submerged. To make the most of snorkel technology and to effectively employ snorkels in the face of the latest advances in anti-submarine

countermeasures, militaries would have to shift their priorities and pursue new clean sheet submarine designs with hulls optimized for submerged operation. This required considerable research and development as well as testing.

Although the aforementioned dynamics would, in time, apply to all states operating submarines, it is nevertheless important to recognize that the above-characterized situation was, all things considered, a uniquely German experience in the Second World War context. Japan, for its part, did not undertake large-scale independent submarine operations in the manner of Germany and the Pacific Ocean was in any event far larger and could not be monitored and protected by surface warships and aircraft in the same manner as the United Kingdom and United States could in the North Atlantic. The British submarine force, which was primarily operating in a supporting role against much smaller German and Italian surface and merchant fleets in the Norwegian Sea, the Baltic, and the Mediterranean, did not encounter a concerted anti-submarine campaign. The American submarine force, which was undertaking a large-scale anti-shipping campaign of its own against Japan, faced weak and largely ineffectual Japanese anti-submarine warfare efforts that were incomparable to the challenges faced by the German submarine fleet in the Atlantic. Germany's predicament in undersea warfare was, as such, unique, and it was, therefore, no accident that Germany, a country at the forefront of many areas of military technology with internationally competitive industries in this timeframe, would undertake pioneering efforts to develop large ocean-going submarines that were optimized for submerged operations as part of a desperate gambit to restore its fast-collapsing submarine-centric naval capabilities.

9.9.2 German Snorkel-Equipped Submarines

Given the complexity of designing a new submarine and the long lead times required to build a submarine, Germany first retrofitted existing Type VII submarines with snorkels and designed new versions of the Type VII that would be built with snorkels as a standard piece of equipment. The Type VII submarine was the German submarine force's workhorse but was a mid-1930s design with limited room for design growth. Moreover, a retrofitted or new-build Type VII equipped with a snorkel still retained a hull design that was optimized for surfaced operations and could not, as such, enjoy the full benefits of a snorkel and, no less importantly, the mere addition of a snorkel to an existing submarine design did not address the enduring limitations of battery technology. For these reasons, Germany – and, in the future, all militaries – required a new clean sheet design in order to take full advantage of the technological opportunities offered by snorkels in a context in which advances in countermeasures meant that submarines could no longer practically operate on the surface as submersibles as they had been standard practice for decades.

In this context, Germany initiated the development of several radically new, that is radically different, submarine designs, of which the larger sea-going Type XXI and the smaller coastal Type XXIII are the most notable. Although these were the first submarines designed to primarily operate submerged to enter frontline service, neither the Type XXI nor the Type XXIII were developed in a vacuum. It is therefore productive to briefly examine the history of efforts toward the development of turbines using hydrogen peroxide, which never entered frontline service, as well as closed-cycle diesel engines during the Interwar Period. In so doing, we will be able to better appreciate the scale of

research and development efforts that were required to simply produce working prototypes of submarines employing radically different propulsion systems.

The designers of submarines were well aware of the limitations imposed on submarines through the use of a diesel engine paired with the combination of electric batteries and an electric motor and periodically pursued breakthroughs in propulsion technology over the Interwar Period. For the most part, these efforts focused on increasing the submerged speed of submarines given that surfaced diesel-powered submarines with hulls optimized for surface operations had many shortcomings, but surfaced speed was not one of them unless a military was adamant that its submarines had to match the speed of largest surface ships, a goal that was increasingly unrealistic as the largest surface ships approached a maximum speed of 30 knots over the Interwar Period. Stated differently, designers experimented with novel propulsion systems not to develop "true" submarines that primarily operated submerged but to develop submersibles that could operate at a higher speed while submerged than electric battery technology allowed.

Diesel engines constituted a compact and, no less importantly, a very efficient propulsion system, and militaries eagerly adopted diesel engines where possible. By the early 1910s, larger ocean-going submarines were almost universally powered by diesel engines and militaries had few complaints about the speed or range of these primarily surfaced submersibles. Once submerged, however, underwater speed-range were severely limited by the poor energy density and specific energy of electrical batteries relative to combustion engines using petroleum products. As a result, designers in several countries experimented with concepts that would allow a diesel engine to operate while the

submarine was submerged. Ordinarily, this was impossible within the enclosed hull of a submerged submarine as a submarine only carried fuel for the diesel engine and not any oxidizer. Instead, submarines required a means of drawing in atmospheric oxygen to facilitate combustion as well as a means of discharging noxious combustion products out of the submarine. If, however, a supply of oxygen could be stored within a submarine, it would be possible to develop what amounted to a closed-cycle diesel engine that could operate while a submarine was submerged.

With stored oxidizer onboard, typically in the form of liquid oxygen, a submarine could operate its efficient diesel engine underwater and thereby operate at greater speed. This concept was appealing in principle, but it was extremely difficult to implement such that, with the singular exception of the Soviet Union's unsuccessful early Cold War Quebec-class coastal submarines, no military has ever productively fielded a submarine employing closed-cycle diesel engine(s) for underwater propulsion. Liquid oxygen had to be safely stored on a submarine and large volumes were required to propel a submerged submarine at a high speed and/or over long distances. In this context, designers searched for alternatives to liquid oxygen and closed-cycle diesel engines, and an enterprising German engineer, Hellmuth Walter, developed the eponymous Walter turbine in the 1930s.

Walter's propulsion system used a catalyst to decompose high-test peroxide, a highly concentrated formulation of hydrogen peroxide, into steam and oxygen. By adding fuel to the mixture, Walter developed a turbine capable of operating without atmospheric oxygen that could propel a submerged submarine to a much higher speed. In 1940, Germany launched the V-80, a very small unarmed experimental submarine employing a

Walter turbine that attained an unprecedented submerged speed of 28 knots at a time when ocean-going submarines struggled to attain 10 knots whilst submerged.[43] On paper, the Walter turbine was both a technological marvel and a major technological breakthrough. In practice, however, it was exceedingly impractical and difficult to scale given the enormous volumes of hydrogen peroxide that would be required to power a much larger submarine with a meaningful endurance and payload that would invariably have to primarily operate on the surface while using a diesel engine to undertake long-range missions (alternatively, a submarine employing a Walter turbine would be restricted to short duration coastal operations). Moreover, hydrogen peroxide, like compressed oxygen in torpedoes and liquid oxygen used in closed-cycle diesel propulsion systems, was difficult to store and the added cost and complexity only offered a submarine the ability to operate at much higher speeds while submerged for a brief time. In every other respect, a submarine equipped with a Walter turbine was no more capable than contemporaneous diesel-powered submarines.

Neither closed-cycle diesel nor Walter turbines constituted a practical propulsion system for a new clean sheet submarine design and these exotic and largely untested propulsion systems offered nothing to address the limitations of existing submarine technology in a manner that would allow the German submarine fleet to recover from its rapidly deteriorating position in the intense wartime measure-countermeasure competition that characterized undersea warfare in the North Atlantic in the ca. 1942-1945 timeframe. Germany was therefore forced to reconceptualize both submarine design and submarine operations and this resulted in the development of two *electroboot* designs

[43] Gudmunur Helgason, "The Walter U-boats," Uboat.net, No Date, https://uboat.net/types/walter_hist.htm

that, as the German name indicates, were characterized by a greater reliance on electric battery technology.

The two *electroboot* designs that were built and employed by the end of the war, the larger ocean-going Type XXI and the smaller coastal Type XXIII, were characterized by streamlined hulls optimized for submerged operation, a snorkel, and a far greater number of battery cells than found on previous submarines. Together, these enabled the development of the first submarines that were capable of primarily operating at snorkel depth and doing so at a higher speed than the submarine could attain whilst surfaced. As a result, these submarines could avoid aerial detection and transit to deployment areas as well as pursue surface ships without being forced to operate at slow speeds in the manner of both converted and new-build Type VII submarines equipped with snorkels. Hence, both the Type XXI and Type XXIII could operate as "true" submarines rather than submersibles that primarily operated on the surface.

Although uncontestably innovative and reflective of a conceptual breakthrough that offered considerable (nominal) gains in fielded military capabilities at a time when successive advances in British and American anti-submarine capabilities had largely neutralized the German submarine force, it is important to recognize that while many *electroboot* submarines were built, few entered service by the end of the war. As a result, the real-world gains in Germany's fielded military capabilities were very limited (i.e., the capability gains nominally engendered by technological change were largely unrealized) to such a degree that the Type XXI and Type XXIII may as well have not existed during the Second World War. Nevertheless, it is important to consider these novel submarines because these offered a realistic possibility of blunting if not negating the latest advances

in anti-submarine capabilities had the state making this technological advance – not necessarily Germany – had greater time and resources to troubleshoot new clean sheet submarine designs. Stated differently, the Type XXI and Type XXIII are important for the present purposes of understanding the military effects of technological change independent of their German origins and, more generally, independent of the highly contingent context and combat dynamics of the Second World War. Of the two pioneering late-war new German submarine designs, the Type XXI was more important as it was capable of long-range operations in open water and more conventional coastal submarines were more difficult to counter even with the latest advances in anti-submarine capabilities. For this reason, I will now turn to the ocean-going Type XXI and disregard the coastal Type XXIII.

The Type XXI was a large submarine by Second World War standards with a displacement of around 1,600 tonnes and therefore was considerably larger than either the Type IX or the Type VII designs that constituted the vast majority of the German submarine fleet throughout the war. Not only did the streamlined hull and snorkel allow the Type XXI to undertake long-range operational patrols at high speeds using a diesel engine at snorkel depth (without operating on the surface), but the greater number of battery cells and the streamlined hull allowed the submarine to operate for several days (at low speed) or undertake high speed dashes while submerged. This allowed the Type XXI to approach, attack, and escape a convoy while submerged. Attesting to this radical reconceptualization of both submarine design and operations, the Type XXI was not equipped with any deck guns and exclusively relied upon twenty-three torpedoes launched through six tubes to attack enemy ships. To facilitate underwater attack as a

primarily underwater naval combatant, the Type XXI was equipped with the most capable passive hydrophone systems then in existence. With snorkels generating a wake that could be detected by enemy ships and aircraft, the Type XXI incorporated a mast-mounted radar warning receiver and a mast-mounted search radar. When altered to the presence of enemy ships and/or aircraft, the Type XXI could dive to a depth of around 200 meters and return to snorkel or periscope depth at a time of its choosing.

Notwithstanding the capability gains that the Type XXI's design offered, it is important to recognize that it was no wonder weapon, not least because most of the 118 submarines nominally completed before the war had a range of production defects that delayed the start of operational deployments.[44] The shortcomings of the Type XXI were not, however, exclusively teething problems associated with the rushed wartime production of a radically new design in an increasingly desperate military and industrial context. The snorkel at the heart of the Type XXI, for example, could be detected by the latest higher frequency/shorter-wavelength radars. While Germany attempted to develop radar-absorbing material to mitigate against this threat, the snorkel also generated a wake, and the submarine could be detected through both active sonar and passive hydrophones as well as other sensors that did not exist in service-ready form by the end of the war (such as an Autolycus, a device carried by aircraft that detected the distinctive exhaust fumes from a diesel engine and could thereby detect the presence of a submarine in the open ocean, or a magnetic anomaly detector, which could detect the ferrous hull of a submerged submarine). Furthermore, the streamlined hull and greater number of battery

[44] This issue has been recognized even in writings that otherwise emphasize the importance of the technological breakthroughs encapsulated in the Type XXI's design. See, for example, Marcus O. Jones, "Innovation for its Own Sake: The Type XXI U-boat," *Naval War College Review* 67, No. 2 (2014): 1-14.

cells offered the Type XXI unparalleled underwater endurance (or speed), but the submarine was just as vulnerable to depth charges and torpedoes as any other submarine and was also vulnerable to gunfire while at snorkel depth. And while the Type XXI was the first ocean-going "true" submarine, it otherwise incorporated much the same components and armaments as the latest versions of Interwar Period origin submarines that Germany built throughout the war. Ultimately, the Type XXI offered the potential of considerable capability gains (or at least a template for such a submarine) without being a wonder weapon and demonstrated that wartime gains in anti-submarine capabilities were reversible but it was ultimately developed too late in the war to have any appreciable impact on fielded military capabilities.

Much like the First World War experience, the Second World War experience with submarine technology provides little evidence that advances in submarine technology engendered sizeable and abrupt shifts in the fielded military capabilities of states. Not only were the cumulative effects of gradual Interwar technological advances limited, but the intense counter-measure competition in undersea warfare that ensued put the final nail in the coffin of the very concept of a submarine as a submersible torpedo boat that defined submarine technology for over half a century. States undoubtedly employed submarines to considerable effect in the Second World War, but this was not a result of technological change and the highly variant experiences of the German submarine force in the Atlantic and the American submarine force in the Pacific serve to highlight that a litany of non-technological confounding factors strongly shaped the effectiveness of submarines.

Chapter 10: Reflecting on the Implications of Advances in Submarine Technology in the ca. 1800 to 1945 Timeframe.

In Chapter Three, I argued that a range of monadic and dyadic factors suggest that technological change is an unsuitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in fielded military capabilities. Not only does technological change tend to be protracted and have considerably lagged effects as concepts are converted into prototypes that undergo testing before being produced and later fielded in militarily significant numbers, but states have considerable scope for adopting, countering, and offsetting new technologies to blunt if not negate the effects of technological change on their military standing. Chapter Three was, however, an exclusively analytical exercise that did not examine real-world cases. Part Two's detailed longitudinal examination of the shifts engendered by advances in submarine technology over the ca. 1800-1945 timeframe provides empirical support for the analytical framework developed in Chapter Three and thereby reinforces the central claim that technological change is not a suitable candidate for a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

In Chapter Five, I explained that submarine technology was subject to nearly a century of protracted evolutionary development during which military submarines were possible but not practical and, importantly, were uncompetitive relative to other military systems. As a result, decades of technological development had negligible effects on the fielded military capabilities of states. While submarine technology experienced what amounted to a sustained growth spurt from the late 1890s onward, technological development began to plateau by the early 1910s as a result of long-running

technological limitations covered in Chapter Six. Compounding matters for then-contemporary states, the effects on fielded military capabilities lagged behind technological change because states did not have infinite resources or the capacity to develop, test, build, and field large numbers of the latest, more capable submarines overnight. In this respect, the growth spurt that submarine technology underwent from ca. 1890-1910 simultaneously rendered submarines a more practical military technology while constraining the effects of technological change on the fielded military capabilities of real-world states operating in a world of finite resources. Given the above considerations, when submarines of varying design vintages were employed in the First World War, not only were submarines far from highly effective in sinking surface ships for purely monadic reasons – for reasons rooted in the limitations of submarine technology – but submarines proved susceptible to basic countermeasures that greatly blunted the effects of advances in submarine technology on fielded military capabilities. Importantly, these countermeasures were far from perfect and did not have to be perfect in order to be effective in blunting if not negating the effects of advances in submarine technology on the fielded military capabilities of states. As a result, and as explained in Chapter Seven, the performance of submarine technology in its first measure-countermeasure competition did not live up to pre-war expectations and advances in submarine technology did not result in a sizeable and abrupt shift in the balance of naval capabilities.

While an argument can be made that submarines were prematurely employed in combat during the First World War – such an argument would of course be supportive of the central claims of this dissertation project – the extension of the longitudinal analysis

into the Interwar Period in Chapter Eight highlighted the enduring limitations of submarine technology, thereby attesting to the regularity of protracted technological development and the lagged effects of technological change on fielded military capabilities. As explained in Chapter Eight, submarine technology gradually approached a state of maturity over the Interwar Period and although states pursued more radical designs and operational concepts for submarines, the enduring limitations of propulsion technology, and the limited scope of developments in other aspects of submarine technology, considerably moderated the effects of technological change on the fielded military capabilities of states independent of the resource-constrained environment that Interwar militaries tended to operate in. Tellingly, the Second World War experience, which unfolded some 140 years after the first submarine underwent testing and some 40 years after the first practical self-propelled submarine underwent testing, further attests to the dynamic of continuity and the gradual accrual of modest gains in military capabilities emanating from technological change. As explained in Chapter Nine, Second World War submarines were, for the most part, late Interwar designs as states simply did not have the time to develop and test and thereafter manufacture and deploy sizeable numbers of radically new submarine designs offering major capability gains. At the same time, the intense wartime measure-countermeasure competition that unfolded over five years rendered the very concept of a submarine as a submersible torpedo (gun) boat increasingly untenable. While late-war German efforts attempted to introduce entirely new submarines to usher in a new era of undersea warfare and thereby give submarines the upper hand in the late-war measure-countermeasure competition, by 1945, as with

prior decades, advances in submarine technology tended not to engender sizeable and abrupt shits in the fielded military capabilities of states.

All things considered, the experience with advances in submarine technology over the ca. 1800-1945 timeframe – 145 years of experience if one starts with the human-powered *Nautilus*, 85 years of experience if one starts with the self-propelled *Plongeur*, and 57 years of experience if one starts with the self-propelled *Gymnote* – undoubtedly offers evidence of the non-zero effects of technological change on fielded military capabilities but does not constitute strong evidence in support of the notion that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

Part Three: Advances in Uncrewed Aircraft Technology ca. 1910-2020

Chapter 11. Uncrewed Aircraft Technology: Setting Up the Analysis

The experience with submarine technology in the ca. 1800-1945 timeframe examined in Part Two of this dissertation highlights the typically protracted nature of technological development and the limited empirical support available for the candidacy of technological change as a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. Advances in submarine technology had non-zero effects on fielded military capabilities, but these typically unfolded gradually and were moderated by both monadic factors and the competitive – dyadic – dynamics of adoption, countering, and offsetting. Part Three of this dissertation will examine whether the more recent and ongoing experience with advances in uncrewed aircraft technology provides support for the candidacy of technological change as a prime mover capable of engendering sizeable and abrupt shifts in fielded military capabilities.

By examining the experience with uncrewed aircraft technology, we can take an additional probe into the generalizability of this dissertation's analytical claims. Furthermore, by examining uncrewed aircraft technology over the ca. 1900-2020 timeframe, we not only have the opportunity to extend the analysis to another physical domain while extending the temporal coverage to ca. 2020, but we also have an opportunity to examine what amounts to a "live case" that international relations scholars have written on in recent years. Uncrewed aircraft technology has, over the past three or so decades, been the focus of considerable discussions and debates that international relations scholars have contributed to,[1] and is increasingly characterized as an area of

[1] For a sample of scholarship on the implications, broadly construed, of advances in uncrewed aircraft technology, see Matthew Fuhrmann and Michael Horowitz, "Droning On: Explaining the Proliferation of Unmanned Aerial Vehicles," *International Organization* 71, No. 2 (2017): 397–418.; Jurgen Altmann and

technological development engendering sizeable and abrupt shifts in fielded military capabilities of consequence for international politics.[2]

Before proceeding to longitudinally examine the history of uncrewed aircraft technology and the effects of technological advances on the fielded military capabilities of states, we must make several analytical determinations. The first analytical determination concerns the establishment of a baseline through which to discern the effects of technological change while the second and third determinations respectively concern the substantive and temporal scope of the analysis. As with the examination of submarine technology in Part Two of this dissertation, these are interrelated issues.

Frank Sauer, "Autonomous Weapon Systems and Strategic Stability," *Survival* 59, No. 5 (2017): 117–142.; Sarah Kreps and Micah Zenko, "The Next Drone Wars: Preparing for Proliferation," *Foreign Affairs* 93, No. 2 (2014): 68–79.; Antonio Calcara, Andrea Gilli, Mauro Gilli, Raffaele, and Ivan Zaccagini. "Will the Drone Always Get Through? Offensive Myths and Defensive Realities." *Security Studies* 31, No. 5 (2022): 791-825.; Antonio Calcara, Andrea Gilli, Mauro Gilli, and Ivan Zaccagini, "Why Drones Have Not Revolutionized War: The Enduring Hider-Finder Competition in Air Warfare," *International Security* 46, No. 4 (2022): 130-171.; Andrea Gilli and Mauro Gilli, "The Diffusion of Drone Warfare? Industrial, Organizational, and Infrastructural Constraints," Security Studies 25, No. 1 (2016): 50–84; Michael Horowitz, Joshua Schwartz, and Matthew Fuhrmann, "Who's Prone to Drone? A Global Time-Series Analysis of Armed Uninhabited Aerial Vehicle Proliferation," *Conflict Management and Peace Science* 39 No. 2 (2022): 119–142.; Andrea Gilli and Mauro Gilli, "The Diffusion of Drone Warfare? Industrial, Organizational, and Infrastructural Constraints," *Security Studies* 25, No. 1 (2016): 50-84.

[2] Consider, for example, Adam Stulberg, "Managing the Unmanned Revolution in the U.S. Air Force," *Orbis* 51, No. 2 (2007): 251–265.; Amy Zegart, "Cheap Fights, Credible Threats: The Future of Armed Drones and Coercion," *Journal of Strategic Studies* 43, No. 1 (2020): 6-46.; Michael Mayer, "The New Killer Drones: Understanding the Strategic Implications of Next Generation Unmanned Combat Aerial Vehicles," *International Affairs* 91, No. 4 (2015): 765-780.; P. W. Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (New York: The Penguin Press, 2009).; Christopher Coker, *Warrior Geeks: How 21st Century Technology is Changing the Way We Fight and Think about War* (London, UK: Hurst, 2013); Peter L. Bergen and Daniel Rothenberg, editors., *Drone Wars: Transforming Conflict, Law, and Policy* (Cambridge, UK: Cambridge University Press, 2014); Hugh Gusterson, *Drone: Remote Control Warfare* (Cambridge, MA: MIT Press, 2016); Michael J. Boyle, *The Drone Age: How Drone Technology Will Change War and Peace* (Oxford, UK: Oxford University Press, 2020).; Sarah E. Kreps, *Drones: What Everyone Needs to Know* (Oxford, UK: Oxford University Press, 2016); Michael P. Kreuzer, *Drones and the Future of Air Warfare: The Evolution of Remotely Piloted Aircraft* (London, UK: Routledge, 2016)

For a notable and relatively early counterargument, consider Jeffrey M. Sullivan, "Evolution of Revolution? The Rise of UAVs," *IEEE Technology and Society Magazine* 25, No. 3 (2006): 43-49.

Given the aim of discerning whether there is empirical support for the notion that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states, I will examine the degree to which technological advances allowed militaries to turn to uncrewed aircraft to *supplant* crewed aircraft in their capability sets. By supplanting, I refer to a process by which technological change allows a military to benefit from capability gains by replacing an existing capability set(s) wholesale. As explained in the prior examination of the implications of advances in submarine technology, the degree to which a new military technology supplants existing technologies in the capability sets fielded by militaries constitutes an important indicator of the overall magnitude of effects engendered by technological change on the fielded military capabilities of states.

Although new technologies can offer militaries the capacity to undertake new roles, the implications for the fielded capabilities of states will necessarily be limited unless and until the new technology can undertake existing roles (or if new roles emerge that supplant existing roles or otherwise make existing roles less relevant). If a new technology does not supplant existing technologies in the undertaking of existing roles, it will likely coexist in a hybrid force structure alongside existing technologies and this will inherently moderate the overall implications of technological change on the fielded military capabilities of states not least because the capability set will only be as effective as the weakest link allows.[3] Whereas monadic considerations, above all the scarcity of resources and the necessity of resource allocation, force militaries to deploy hybrid force

[3] While new military technologies tend to coexist alongside existing technologies until the former can be produced and fielded in such numbers as to replace the latter, the dynamic being highlighted here is not driven by the scarcity of resources but the unsuitability of a new technology to wholly replace an existing technology.

structures, the inability of a new technology to supplant existing technologies constitutes a distinctly technological factor contributing to moderated effects on fielded military capabilities.

It is important to recognize that this emphasis on the degree to which uncrewed aircraft have supplanted crewed aircraft is not theoretical or crafted in support of the narrow objectives of this dissertation project; it derives in large part from the extensive speculation and debate as to whether there is a future for crewed combat aircraft in the face of progressive advances in uncrewed aircraft technology. From circa 2010 onwards, an increasingly prominent argument has been that uncrewed aircraft constitute the future of aerial combat capabilities. In one particularly notable instance in 2015, a senior American official went as far as to state that the F-35, a crewed fighter aircraft that had not even entered frontline service at the time, would be the United States Navy's "last manned strike fighter."[4] More recently, leading militaries have announced plans to develop and build an increasingly diverse array of uncrewed aircraft (alongside a new generation of crewed combat aircraft).[5] If such claims are accurate and if such announcements are reflective of the trajectory of technology and capability development, then uncrewed aircraft technology may be an excellent place to find evidence that technological advances engender sizeable and abrupt shifts in the fielded military capabilities of states.

---

[4] Sam Lagrone, "Mabus: F-35 Will Be 'Last Manned Strike Fighter' the Navy, Marines 'Will Ever Buy or Fly'," *USNI News*, April 15, 2015, https://news.usni.org/2015/04/15/mabus-f-35c-will-be-last-manned-strike-fighter-the-navy-marines-will-ever-buy-or-fly

[5] Stephen Losey, "US Air Force eyes fleet of 1,000 drone wingmen as planning accelerates," *Defense News*, March 8, 2023, https://www.defensenews.com/air/2023/03/08/us-air-force-eyes-fleet-of-1000-drone-wingmen-as-planning-accelerates/;"U.S. Air Force Presses for Collaborative Combat Aircraft Cost Sharing" *Aviation Week*, February 20, 2024, https://aviationweek.com/defense-space/aircraft-propulsion/us-air-force-presses-collaborative-combat-aircraft-cost-sharing/

Before proceeding, it is important to define what is and is not an uncrewed aircraft and thereby delineate the scope of Part Three of this dissertation project. This dissertation espouses a broad conception of uncrewed aircraft technology, one that encompasses any form of automated human-made object that is (A) capable of flight; (B) does not carry any human operators onboard; and (C) can undertake some form of deliberate maneuvering (i.e., the object must not maneuver after entering flight solely as a result of the interplay of gravitational attraction and atmospheric effects). To undertake some form of deliberate maneuvering while in flight, uncrewed aircraft, which by definition do not have a human operator onboard, require *some level* of automation. Automation can be minimal, as in guided munitions of the manual command line of sight (MCLOS) variety that lacks an autopilot and is wholly dependent on the command inputs made by a remote human operator (i.e., a human in the loop) that controls the uncrewed aircraft in flight through some communications medium. Automation can be much more comprehensive such that an uncrewed aircraft can operate without a human in the loop whether by following a preprogrammed flight path or, in a far more sophisticated form, by autonomously responding to the environment as observed through onboard sensors. Automation can be and increasingly tends to be, somewhere in between such that an uncrewed aircraft typically has some form of autopilot but requires human supervision and responds to new commands issued by a remote human operator. Stated differently, for the present purposes, uncrewed aircraft technology is nothing more than the manifestation of *some level* of automation in military aircraft technology. By this standard, a passenger aircraft or a (crewed) combat aircraft equipped with an autopilot are also examples of partial – and very limited – automation that is unsuitable for many tasks

and which therefore requires human supervision if not regular human intervention (these are not, however, examples of uncrewed aircraft because said aircraft cannot operate in an unmodified state without a human pilot onboard). Stated differently, automation is best conceptualized as existing on a spectrum that encompasses both crewed and uncrewed military aircraft.

Although this dissertation project espouses a broad conception of uncrewed aircraft technology, this chapter does not examine every type or category of uncrewed aircraft ever conceived, tested, built, or put into use. Instead, this chapter focuses on a specific subset of uncrewed aircraft technology, a subset that has increasingly been at the center of debates about the implications of technological change over the past twenty or so years. Specifically, this chapter examines the implications of advances in uncrewed aircraft that are (A) large – have a takeoff weight of over 100 kilograms; (B) designed to be reusable; (C) fixed-wing aircraft; (D) do not – and cannot – carry a human pilot onboard; (E) intended for roles other than training and testing support purposes. Stated differently, this chapter focuses on the capability gains engendered by advances in *large reusable fixed-wing uncrewed aircraft technology* (hereafter referred to as "LRFW uncrewed aircraft").

Given the present focus on the degree to which advances in uncrewed aircraft technology have allowed uncrewed aircraft to supplant crewed aircraft, Condition A – large, weighing over 100 kilograms – reflects upon the often-overlooked consideration that an uncrewed aircraft design must have an adequate range (i.e., onboard fuel) and payload (for sensors and/or munitions) to effectively supplant crewed combat aircraft in the latter's existing roles. It is, therefore, important to avoid the erroneous and

widespread comparison of highly dissimilar aircraft and instead engage in what may be characterized as "citrus-to-citrus" comparisons of crewed and uncrewed aircraft. Just as it is analytically pointless to compare a tank equipped with a large caliber gun with a handheld assault rifle, it is analytically pointless to compare a 15,000-kilogram crewed aircraft with a (fully loaded) combat range of 600 kilometers, a maximum speed of 1,500 km/h, and a payload of 6,000 kilograms with a 1.5-kilogram uncrewed aircraft that has a maximum range of five kilometers and a maximum speed of 50 km/h. Whatever capability gains offered by uncrewed multirotors such as quadcopters and small – under 100 kilogram – fixed-wing uncrewed aircraft designs, these cannot offer remotely comparable performance in terms of range, speed, and payload and cannot, therefore, supplant the vast majority of crewed military aircraft designs in their primary roles even if such smaller genera of uncrewed aircraft technology affect the fielded military capabilities of states in other ways.

Condition B – reusability – is intended to make a distinction between uncrewed aircraft that are designed to be reusable, in the manner of crewed military aircraft, and those that are primarily, if not exclusively, intended to function as single-use munitions, whether surface-to-air, air-to-surface, air-to-air, or surface-to-surface.[6]

Condition C – fixed-wing aircraft – excludes rotary-wing aircraft including multirotors such as quadcopters as well as lighter-than-air aircraft irrespective of their size and thereby facilities like-for-like comparisons.

[6] It is worth explicating that condition B (reusability) excludes prominent specimens of uncrewed aircraft technology from this analysis irrespective of the level of automation and irrespective of the means of propulsion. Specifically, it excludes both the likes of a propeller-driven fixed-wing Iranian Shahed-136, which flies on a preprogrammed flight path, and an American turbofan-powered fixed-wing Tomahawk cruise missile, which also flies on a preprogrammed flight path but is equipped with a satellite datalink to enable in-flight retargeting.

Condition D – no onboard human pilot – excludes the least automated aircraft designs, which are typically those in which automation exists to augment rather than substitute for – replace – the onboard pilot(s) (i.e., a crewed aircraft, civilian or military, with an autopilot designed to augment the pilot by reducing pilot workload).

Condition E is intended to exclude uncrewed aircraft that, while potentially reusable, are primarily designed to serve as aerial targets for training and testing purposes. Such uncrewed aircraft tend to be very limited in terms of automation, payload, and range befitting their primary role as affordable and expendable targets.

Given the tenor of the prevailing discourse on the implications of advances in uncrewed aircraft technology,[7] it is productive to anticipate potential critiques and explain why the substantive focus on large reusable fixed-wing uncrewed aircraft – LRFW uncrewed aircraft – is justified. Absent a fundamental breakthrough in the specific energy and energy density of the fuels that can be practically used[8] by aircraft powered by a combustion process – or a fundamental breakthrough in the specific energy and energy density of electric batteries, which remain woefully inferior to petroleum products – a crewed aircraft (with up to two onboard crew members, as has become customary for over fifty or so years) and an uncrewed aircraft with equal performance in terms of speed, range, payload, rate of climb, maneuverability will, all things equal, be of broadly

[7] Consider, for example, "Killer Drones Pioneered in Ukraine Are the Weapons of the Future," *The Economist*, February 8, 2024, https://www.economist.com/leaders/2024/02/08/killer-drones-pioneered-in-ukraine-are-the-weapons-of-the-future; "How Cheap Drones Are Transforming War in Ukraine," *The Economist*, February 5, 2024, https://www.economist.com/interactive/science-and-technology/2024/02/05/cheap-racing-drones-offer-precision-warfare-at-scale

[8] Hydrogen has a specific energy – energy "density" in terms of mass – greater than any petroleum product but its low density at ambient temperatures is such that it has a much poorer energy density – in terms of volume – in practice. Notwithstanding ongoing research that may result in future technological breakthroughs, hydrogen remains an impractical fuel source for aviation as of this writing.

comparable size and weight (i.e., density, assuming the use of comparable materials and production methods in their construction).

While an uncrewed aircraft can dispense with the enclosed volume and weight of structural material required to house the onboard crew and any environmental support equipment required for the onboard human crew, an uncrewed aircraft will nevertheless need to retain some of the nominally saved weight and volume to house computers for automated flight controls, sensors, and datalinks to substitute for an onboard pilot irrespective of whether the uncrewed aircraft is fully automated or requires a (remote) human in the loop to operate. All things equal, the larger and heavier the aircraft (i.e., the greater the range and payload of the aircraft), the smaller the proportion of the total weight and volume that will be saved by substituting an onboard human crew with onboard computers, sensors, and datalinks. Given that range (i.e. fuel capacity) and payload are two of the primary determinants of the roles for which an aircraft, whether crewed or uncrewed, is suitable, and two of the primary contributors to airframe weight, the primary differentiator of an uncrewed aircraft design from a crewed aircraft design is the extent to which it can match or exceed the performance of, and as such the capabilities offered by, an onboard human pilot in one or more roles.

Dispensing with onboard human operators does, however, have an implication that is important to highlight even though it primarily applies to the types of uncrewed aircraft that are beyond the scope of this dissertation (i.e., uncrewed aircraft that are not large and/or not reusable). Stated plainly, dispensing with the weight and volume taken up by onboard human operators and any environment support systems allows for the *miniaturization* of aircraft design such that an uncrewed aircraft can be smaller and

lighter than a human operator. This is, by definition, impossible with any crewed aircraft design, and miniaturization consequently opens a range of possibilities for aircraft designers and militaries much as it does with uncrewed ground, surface marine, and subsurface marine vehicles. Consequential though this dynamic of miniaturization may be, it is important to appreciate that there is nothing new about miniaturized uncrewed aircraft. As I will explain in the following pages, since the 1940s, if not earlier, militaries have adopted *some level* of automation to enable the development and use of miniaturized uncrewed aircraft, broadly defined. This applies to everything from cruise missiles to much smaller anti-tank missiles as well as unpowered (i.e., glide) air-to-surface munitions. Miniaturization is not, as such, a recent phenomenon and, through 2020, miniaturization has primarily allowed militaries to augment and complement crewed combat aircraft in the form of guided munitions, including both air-launched guided munitions capable of undertaking so-called standoff engagements that reduce the exposure of onboard human operators to air defenses, as well as highly automated guided munitions such as cruise missiles. With the present objective being an assessment of technological change in terms of the degree to which advances in uncrewed aircraft technology have allowed militaries to supplant crewed aircraft with uncrewed aircraft, we can productively limit the scope of this analysis to large, reusable fixed-wing uncrewed aircraft and thereby exclude smaller miniaturized uncrewed aircraft (i.e., aircraft that weigh under 100 kilograms, a grouping which includes all human-portable and handheld uncrewed aircraft designs).

Having established an analytical baseline and the substantive scope of this chapter, we are left with the task of defining the temporal scope of analysis. As explained

throughout this dissertation, it is imperative to undertake longitudinal assessments of the implications of technological change and not simply begin analysis at a timeframe in which, with the assistance of hindsight, a technological advance is deemed to be consequential. This case study will therefore comprise a longitudinal assessment of the implications of advances in uncrewed aircraft technology in the ca. 1910-2020 timeframe. As I will explain in the following pages, the history of uncrewed aircraft technology is over one hundred years old as of this writing even if the history of practical and effectual uncrewed aircraft technology is considerably shorter. As with the experiences with submarine technology examined in Part Two of this dissertation, the history of advances in uncrewed aircraft technology encompasses long periods of negligible and non-existent effects on the fielded military capabilities that are part and parcel of the historical record and should therefore inform our understanding of pace and magnitude of the shifts that technological advances engender in the fielded military capabilities of states.

Before proceeding, it is productive to explicate that Part Three of this dissertation focuses solely on the military effects of advances in uncrewed aircraft technology in order to determine whether technological change in this area supports the proposition that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. Part Three does not, as such, examine the wider implications of advances in uncrewed aircraft technology including the political, diplomatic, legal, and ethical implications that much of recent international relations scholarship examining uncrewed aircraft technology has tended to focus on. More generally, this dissertation does not contend that advances in uncrewed

aircraft technology have had zero effects on the fielded military capabilities of states rather, the focus is limited to the pace and change of shifts engendered in the fielded military capabilities of states in both monadic terms and in dyadic terms – in relation to other states. For this reason, Part Three of this reason will not examine the military effects of uncrewed aircraft technology in counterinsurgency, counterterrorism, border security, and similar operations directed against primarily non-state adversaries with weak if not non-existent air defenses.

Chapter 12: Uncrewed Aircraft Technology ca. 1910-1945

Having established an analytical baseline, a starting point, and a definition of what a LRFW uncrewed aircraft is, we can now examine whether advances in uncrewed aircraft technology in the ca. 1910-1945 timeframe engendered sizeable and abrupt shifts in the fielded military capabilities of states. If technological change is a suitable candidate for the role of a prime mover engendering shifts in the fielded military capabilities of states that are both sizeable and abrupt, then we should be able to find evidence of this in the experience with advances in uncrewed aircraft technology across time and space. Although the ca. 1910-1945 timeframe examined in this chapter starts much earlier than most writings on uncrewed aircraft technology, it is problematic to select cases or select the temporal scope of a study on the basis of the outcome of interest. As such, if there is evidence that advances in uncrewed aircraft technology in the ca. 1910-1945 timeframe engendered sizeable and abrupt shifts (gains) in the fielded military capabilities of adopting states, then this will be evidence in support of the proposition that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. If such evidence is not to be found, however, then any evidence in support of the proposition will be weakened absent supporting evidence from later timeframes that will be examined in the subsequent chapters of this longitudinal analysis.

Writings on the implications of advances in uncrewed aircraft technology tend to begin their analyses around the mid-1970s, which is when a new generation of smaller and lighter – miniaturized – uncrewed aircraft were developed and fielded. With the benefit of hindsight, it is possible to discern that the technological advances made in this

timeframe, whilst limited in terms of the (immediate) effects engendered on the fielded military capabilities of states at the time, put into motion technological trends that, after a protracted gestation period, engendered progressively greater effects on fielded military capabilities from the mid-1990s onward. The technological advances made in the 1970s and 1980s did not, of course, emerge in a vacuum and, more importantly, are best understood as the emergence of *new forms* of uncrewed aircraft made possible by technological advances rather than the beginnings, let alone the entire future, of uncrewed aircraft technology. Just as the history of submarine technology is characterized by decades of protracted developmental efforts in which gradual technological advances had negligible if not non-existent effects on the fielded military capabilities of states, including decades in which a military submarine was *possible* but not practical, the history of uncrewed aircraft technology includes decades of technological advances that did not amount to much in terms of effects on fielded military capabilities.

12.1 The First Uncrewed Fixed-Wing Aircraft

The history of human-crewed aircraft capable of powered and controlled flight (with fixed-wing aircraft) begins, of course, with the Wright Brothers in 1903. Less well-known is the fact that the history of uncrewed fixed-wing aircraft technology began within a decade of that milestone when crewed fixed-wing aircraft were equipped with gyroscopes.[9] Making use of gyroscopes initially developed for surface ships, designers developed gyroscopic autopilots that allowed a crewed aircraft to maintain level flight on

[9] For this, the designer, Elmer Sperry, was awarded the Collier Trophy in 1914.
For an excellent overview of early efforts aimed at developing uncrewed vehicles of all types, including aircraft, see H.R. Everett, *Unmanned Systems of World Wars I and II* (Cambridge, MA: MIT Press, 2015)

a fixed bearing, thereby reducing pilot workload and fatigue. At a time in which pilots had to physically contest with the controls for command over an aircraft, the advent of autopilots allowed, in principle, a pilot to undertake other tasks, such as checking a map. As such, the history of uncrewed aircraft technology begins with a means of augmenting the human pilot onboard rather than substituting for the pilot as the level of automation was exceedingly limited when placed on a spectrum ranging from no automation to full automation. The military implications of this relatively minor technological development, one overshadowed by successive incremental advances in (crewed) aircraft technology in the years leading up to the First World War, were negligible at the time but nevertheless laid the foundations for more consequential future developments in uncrewed aircraft technology.

Wartime demands, above all the challenges imposed by the stalemate that characterized the Western Front, led militaries to fund research and development efforts aimed at finding some idea or another that could plausibly offer an improvement in their combat capabilities. This surge of wartime investment led to two efforts that form part of the early developmental history of uncrewed aircraft technology. The first concerns the attempted development in the United Kingdom of uncrewed propeller-driven target aircraft operated remotely by a pilot on the ground by means of a radio. The advent of practical radiotelegraphy in the 1890s resulted in widespread efforts to develop remote-controlled vehicles of all types, including a particularly high-profile demonstration of a radio-controlled boat by Nikola Tesla in 1898.[10] The very development of gyroscopic autopilots in the United States had been motivated by a desire to develop a radio-

[10] "Remote Control," Public Broadcasting Service, No Date, https://www.pbs.org/tesla/ins/lab_remotec.html

controlled aircraft as an onboard autopilot, however crude, offered the potential of reducing the workload of the remote pilot and the risk of a remotely controlled aircraft crashing.[11] Viewed in this light, the wartime British effort to develop a remotely piloted uncrewed aircraft was hardly without precedent. While an uncrewed target aircraft could have served as a useful training tool at a time when militaries needed to train ever-increasing numbers of personnel to provide air defense over the battlefield and at sea against progressively faster, higher flying, and more maneuverable crewed fixed-wing aircraft, the testing of the prototype radio-controlled target aircraft was unsuccessful, and this attempt at developing one of the first fixed-wing uncrewed aircraft had no bearing on the fielded military capabilities of states.

The second wartime developmental effort worth considering – I will not examine a similar third contemporary effort in the interest of brevity[12] – resulted in the Kettering Bug, an uncrewed propeller-driven fixed-wing aircraft designed to serve in a single-use (self) destructive manner that is typically associated with later cruise missiles. Whereas a First World War era radio-controlled aircraft, had one been successfully developed, would have required near-constant human supervision and control inputs (i.e., a human in the loop), the American Kettering Bug relied upon a basic gyroscopic autopilot to maintain a constant heading toward its intended target. After a preset number of engine revolutions, a proxy for the distance to the intended target, a revolution counter would shut off the engine, and simple automated controls would place the aircraft in a dive,

[11] "A Brief History of Drones," Imperial War Museum, No Date, https://www.iwm.org.uk/history/a-brief-history-of-drones#:~:text=The%20Aerial%20Target%2C%20a%20British,was%20tested%20in%20March%201917

[12] The third example is the American Hewitt-Sperry Automatic Airplane, which was conceptually similar to the Kettering Bug. This developmental effort was also unsuccessful and had no effects on fielded military capabilities.

thereby directing the airframe and the 80 or so kilograms of onboard explosives into the ground.[13] The Kettering Bug had a testing record that left much to be desired, few were built, and none were used in combat.[14] Nevertheless, the Kettering Bug, as with the British efforts toward developing radio-controlled uncrewed aircraft, demonstrated what was – and was not – possible and practical at the time in the same manner that early efforts at submarines covered in Part Two of this dissertation, ineffectual though these were, reflected how even unsuccessful attempts technological advances could expand the technological frontier.

These First World War efforts toward developing uncrewed military aircraft were innovative but largely impractical given the immaturity of the core component technologies. While a rudimentary type of aircraft automation existed in the form of gyroscopic autopilots, these were more useful in augmenting onboard human pilots than serving as the sole navigation and guidance system through which an uncrewed aircraft would operate. The state of autopilot technology was not, in other words, capable of directing an uncrewed aircraft toward, let alone from, a specific area, and there was no practical method of controlling what the uncrewed aircraft would do once it reached the target area. With wet film cameras, no onboard navigation system – and no human operator – to determine the location of the aircraft, and no means of sharing imagery captured on film while in flight, early specimens of uncrewed aircraft technology were

[13] "Kettering Aerial Torpedo 'Bug'," National Museum of the United States Air Force, No Date, https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/198095/kettering-aerial-torpedo-bug/

[14] "Gyroscope, Autopilot, Kettering Liberty Eagle 'Bug'," National Air and Space Museum, No Date, https://airandspace.si.edu/collection-objects/gyroscope-autopilot-kettering-liberty-eagle-bug/nasm_A20010524000

wholly unsuited to the types of intelligence, surveillance, reconnaissance (ISR)[15] roles that later LRFW uncrewed aircraft would become closely associated with. Furthermore, without a human operator on board, or a practical means of both controlling a remotely piloted uncrewed aircraft over a radio link *and* operating onboard machine guns to fire small kinetic (i.e., non-explosive) projectiles at progressively faster and more maneuverable fixed-wing aircraft, uncrewed aircraft technology of this era was also wholly unsuited for any anti-aircraft role.[16]

For these reasons, the only conceivable roles for an uncrewed fixed-wing aircraft in the First World War timeframe were, to use later terminology, a cruise missile analog along the lines of the Kettering Bug, or a training target. Without the assistance of any onboard terminal guidance or the bounded navigation system and ballistic computer found inside the head of a trained human pilot, uncrewed aircraft designs like the Kettering Bug were an expensive means of inaccurate and low explosive yield – and thus likely ineffectual – aerial bombardment and were no less vulnerable to air defenses than contemporary crewed aircraft undertaking bombing missions. As a result, these innovative and ground-breaking technological efforts had negligible if not non-existent implications for fielded military capabilities in the 1914-1918 timeframe.

---

[15] The usage of the concept of intelligence, surveillance, reconnaissance – ISR – is anachronistic when referring to the first half of the twentieth century and largely reflects upon late twentieth century technological and doctrinal developments that allowed for military aircraft, whether crewed or uncrewed, to simultaneously undertake multiple roles. Furthermore, the distinction between intelligence, surveillance, and reconnaissance primarily emanates from the discourse that informs military doctrine in the United States. Historically, ISR aircraft were typically referred to as observation or reconnaissance aircraft, but this has changed in recent decades. Nevertheless, in the interest of brevity, I will use ISR instead of repeatedly alternating between intelligence (gathering), reconnaissance, and surveillance (which, unlike reconnaissance, connotates some degree of persistent observation of a given area).

[16] In principle, an uncrewed aircraft could have been used in an anti-balloon (i.e., "balloon busting") role but there is no evidence of any such undertaking and anti-balloon efforts appear to have exclusively involved crewed combat aircraft. This is not surprising given the primitive state of automaton, sensor, and communications technology.

12.2 Technological Developments in the Interwar Period

The end of the First World War was followed by a sharp decline in investment toward the development and manufacture of new military technologies that, for the most part, extended into the mid-1930s. Although the development of uncrewed aircraft was not a priority in any country during the Interwar Period, substantive, if gradually unfolding, advances in a core component technology area – communications technology – progressively expanded the realm of possibilities.

While radio technology existed before the First World War, wartime research and development efforts laid the foundations for major advances in radio technology in the Interwar Period. Primarily undertaken by civilians for civilian applications, technological advances made radios increasingly practical, reliable, and affordable such that relatively compact radios suitable for mobile applications quickly became standard on (crewed) military aircraft and other military systems including submarines and, in time, tanks. Advances in radiotelephony, which enabled commercial radio broadcasting; the development of television broadcasting; and the use of progressively higher frequency ranges of the electromagnetic spectrum, among other developments, created an environment in which militaries and designers gave greater consideration to the practicalities of radio-controlled – remotely piloted – uncrewed aircraft.

Whatever the potential capability gains offered by an uncrewed aircraft design, it is important to appreciate that sustained advances in (crewed) aircraft technology over the Interwar Period meant that proposals for an uncrewed aircraft had to offer some quality that increasingly capable crewed aircraft could not match or exceed. Concerns about the prospect of the aerial bombardment of cities and the threat posed to surface ships by

aircraft equipped with bombs and torpedoes drove demand for a target aircraft against which the operators of air defense guns could train. While relatively small unpowered uncrewed aircraft could be released by a glider tug or launched from a catapult, the inability of such airframes to undertake controlled maneuvers in flight made these low-fidelity surrogates for the progressively faster, more maneuverable, and higher-flying crewed combat aircraft that air defenses would encounter in combat conditions. Although crewed aircraft were by definition excellent substitutes for other crewed aircraft in the role of a target, it was not practical to fire live ordnance at crewed aircraft used in a training support role. As a result, concerns about air defense capabilities on land and at sea created an opening for an uncrewed aircraft that could be employed in a role, however limited, that increasingly capable crewed aircraft could not be practically used for.

By the mid-1930s, several militaries were engaged in research and development efforts aimed at producing a radio-controlled uncrewed target aircraft. Attesting to the engineering challenges that existed even in the context of the aforementioned background advances in radio technology, these developmental efforts did not always succeed. By 1935, however, the British navy developed a functional radio-controlled target aircraft, the DH-82B Queen Bee.[17] Derived from similarly sized crewed aircraft, the Queen Bee's performance in early training exercises, in which it continued flying unscathed, highlighted the deficiencies of naval air defenses and spurred the development of improved gunnery practices, (analog) fire control computers and, in time, the development and application of naval radars for range finding and target tracking.

[17] "De Havilland DH82B Queen Bee," de Havilland Aircraft Museum, No Date, https://www.dehavillandmuseum.co.uk/aircraft/de-havilland-dh82b-queen-bee/

The Queen Bee, the first functional radio-controlled large fixed-wing uncrewed aircraft, constituted a breakthrough in uncrewed aircraft technology but it is important to appreciate its limitations and that of other uncrewed target aircraft developed in the years leading up to the Second World War. These aircraft were stabilized in flight by basic autopilots and maneuvered on the basis of control inputs transmitted by a remote pilot over radio. The control arrangement was limited as the remote pilot had no feedback from the vantage point of the aircraft and had limited fidelity over the controls on the basis of its visual observation, something reflected in the fact that the Queen Bee, a biplane design, was exclusively controlled through manipulation of the throttle, rudder (i.e., yaw control), and elevators (i.e., pitch control), and not the ailerons (i.e., roll control). Such an uncrewed aircraft could not be remotely piloted beyond the line-of-sight of the remote pilot on the Earth's surface and, absent a breakthrough in communications technology, could not provide feedback in the form of transmitted real-time imagery – from the uncrewed aircraft's vantage point – that could be used for navigation or any targeting onboard sensors and/or munitions. Compounding matters, the limitations of then-contemporary radio technology, affected as it was by atmospheric conditions and interference caused by the proliferation of civilian and military radios, meant that remote control operation did not permit the full replication of the performance characteristics of the crewed aircraft designs from which remotely piloted target aircraft were typically derived or converted.

12.3 Technological Developments During the Second World War

By the start of the Second World War, technological advances barely allowed for functional radio-controlled uncrewed target aircraft, let alone more complex designs capable of automated and, as such, unsupervised operation. Several decades of technological development consequently had negligible effects on the fielded military capabilities of states and uncrewed aircraft technology was in no position to supplant crewed aircraft technology in any role other than serving as a target for air defenses to train against. Nevertheless, the technological advances made in the thirty or so years from ca. 1910 through 1940 helped to lay the foundations for the development of guided air-to-ground munitions, one of the first practical and effectual applications of uncrewed aircraft technology toward frontline, as opposed to supporting (i.e., training), military roles.

As with the First World War, wartime necessity resulted in unprecedented levels of investment toward research and development efforts aimed at turning what had previously been far-fetched ideas into reality. While Germany, the United Kingdom, and the United States undertook distinct approaches to aerial bombing,[18] these countries nevertheless quickly encountered a shared reality that advances in aircraft technology (i.e., faster, and more effective fighters and interceptors) and advances in air defense technology (i.e., air defense artillery and radar) over the Interwar Period made aerial bombing an increasingly hazardous undertaking. Compounding matters, unsustainably

[18] Tami Davis Biddle, *Rhetoric and Reality in Air Warfare: The Evolution of British and American Ideas About Strategic Bombing, 1914-1945* (Princeton, NJ: Princeton University Press, 2002). On civilian targeting, see Alexander B. Downes, *Targeting Civilians in War* (Ithaca, NY: Cornell University Press, 2012). On bombing as a means of coercion, see Robert A. Pape, *Bombing to Win: Air War and Coercion in War* (Ithaca, NY: Cornell University Press, 1996)

high attrition rates for aircraft on bombing missions were often not compensated for by the destruction of the intended targets. Aerial bombing against defended targets – as opposed to the empty fields of a training range – tended to be very inaccurate, not least in the context of the highly variable and regularly overcast atmospheric conditions of northwestern Europe, and higher altitude bombing and nighttime bombing, the two primary approaches to mitigate against the threat posed by air defenses, further diminished bombing accuracy. In this context, Germany, the United Kingdom, and the United States, the three countries that pursued the primary aerial bombing campaigns of the war, searched for ways to improve the effectiveness of their (crewed) bomber aircraft. What resulted were multifaceted efforts aimed at overcoming air defenses, improving navigation, and, among other things, the development of guided air-to-ground munitions.

The advances in radio control and autopilot technology made throughout the Interwar Period offered a foundation for designers to experiment with the development of guided air-to-ground munitions. The promise of guided air-to-ground munitions was straightforward: air-dropped bombs were subject to the interplay of gravitational attraction and atmospheric effects and bombardiers had no control over bombs once these were released from the launch aircraft. While careful navigation and bomb sights ostensibly improved bombing accuracy, these approaches remained inadequate in combat conditions as there existed no means of terminal guidance over the individually inaccurate bombs in uncontrolled flight. The challenge was to find a means of controlling the flight path of a bomb (i.e., remotely flying a bomb) as it descended toward the ground without having a human pilot onboard (as in a kamikaze aircraft, which "solved" this

technical problem of navigation and guidance by accepting the death of a human pilot turned single use onboard navigation and guidance system).

To address this challenge, designers turned to Interwar advances in autopilot and radio technology. To be clear, Second World War era designers did not have the luxury of simply opening a proverbial tap from which production and combat-ready designs would immediately flow. The technological advances made in the Interwar Period provided foundations that wartime designers could build upon but considerable research and development efforts, as well as extensive testing, were required before practical guided munitions could be produced and fielded in militarily significant numbers.

Although the United Kingdom was responsible for its substantial number of wartime technological developments, it was Germany and the United States that led the way in the development of guided air-to-ground munitions. By 1943, after several years of development and testing, Germany introduced the Fritz X,[19] and in 1944 the United States introduced the AZON (AZimuth ONly). The Fritz X and the AZON constitute some of the earliest guided air-to-ground munitions – guided bombs – and, more importantly for the present purposes, constitute the first practical uncrewed aircraft designs to enter service that were not target aircraft intended for training purposes. These German and American guided bombs were broadly similar in terms of design and function. Both were uncrewed, albeit minimally automated, unpowered aircraft that a remote human operator – located in a nearby aircraft that released the guided bomb – controlled via radio through the manipulation of a control stick that actuated control surfaces mounted on the bomb. While these guided bombs did not have a robust autopilot

[19] The Fritz X is the most common name given to this German guided bomb. "Bomb, Guided, Fritz X (X-1)," Smithsonian, No Date, https://www.si.edu/object/bomb-guided-fritz-x-x-1%3Anasm_A19840794000

capable of independent navigation toward the intended target, the human operator could rely upon an onboard gyroscope to keep the bomb level in flight. Given the low level of automation possible at this time, the Fritz X and AZON were manual command line of sight (MCLOS) designs in which a human was constantly in the loop as the remote human operator was, in effect, flying the uncrewed aircraft over a radio link as if the bomb was something in the vein of an unpowered radio-controlled hobby aircraft launched from altitude. While the Fritz X and AZON were small, unpowered uncrewed aircraft with a short glide range, Germany and the United States also developed more "aircraft-like" powered air-to-ground munitions. The same combination of radio technology and autopilot technology (i.e., autonomy in flight, however limited) enabled the development of the German Hs 293, a radio-guided rocket-powered glide bomb, and the even more sophisticated (in terms of the guidance system) propeller-driven (i.e., plane-like) American InterState TDR, which introduced television guidance that offered the remote operator feedback in the form of a TV camera feed from the uncrewed aircraft's perspective.

The unprecedented intensity of wartime research and development activity, including the large-scale testing of prototypes, facilitated the pursuit of a wide range of potentially promising uncrewed aircraft designs and component technologies. By the end of the war, Germany was on its way to adapting radio-controlled MCLOS guidance technology to radio-guided rocket-powered guided anti-aircraft munitions (i.e., surface-to-air missiles), an area of technological development in which the United States and the United Kingdom were also active. These developmental efforts did not result in a combat-ready system by the end of the war and were beleaguered by limitations

emanating from technological immaturity but can nevertheless be viewed as meaningful attempts to develop uncrewed substitutes for the likes of the German rocket-powered fixed-wing Me 163 crewed interceptor aircraft and thereby develop a more practical and more effective application of advances in rocket propulsion technology to the role of air defense. At the same time, Germany experimented with wire guidance systems, by which an uncrewed aircraft received commands electrically transmitted through a trailing copper wire, in order to address the inherent vulnerabilities of radio command links to jamming. By so doing, Germany led the way in efforts toward developing wire-guided air-to-air missiles and wire-guided surface-to-surface (i.e., anti-tank) missiles, two other new manifestations of uncrewed aircraft technology. The United States, for its part, also pioneered many advances in uncrewed aircraft technology during the Second World War. Beyond the television guidance systems used on the TDR, perhaps the most impressive of these developmental efforts resulted in the ASM-N-2 Bat, an active radar-homing guided glide bomb that could independently fly to a target detected by the (bomb's) onboard radar for which there existed no German counterpart.[20] The United States also developed the LBD-1 Gargoyle, a radio-controlled rocket-powered anti-ship munition, and the VB-6, a guided bomb with an onboard infrared seeker that allowed for the dispensing of radio guidance as the bomb could independently home in on a suitable heat source without human intervention.

While these are important datapoints in the history of the development of uncrewed aircraft technology, broadly defined, it is nevertheless important to recognize that most of these innovative designs saw limited, if any, combat use during the war, and

[20] "Bureau of Standards SWOD MK 9/ASM-N-2 Bat," Designation Systems, No Date, https://www.designation-systems.net/dusrm/app1/asm-n-2.htmlAS

the results from testing and/or combat use often left much to be desired.[21] Radio-controlled MCLOS munitions, for example, which constituted the most common approach to employing guided munitions at this time, were only as accurate as the vantage point and dexterity of the remote human pilot allowed. Gyroscopic autopilots assisted the pilot in the remote operation of the uncrewed aircraft and television guidance, as used on the American TDR, could provide feedback from the aircraft's vantage point so that the remote pilot could remotely operate the aircraft beyond their eyesight, but even television guidance still meant that a human was required in the loop for the entire flight such that the uncrewed aircraft could not operate autonomously. Moreover, the limitations of radio technology meant that the remote pilot had to remain no more than several kilometers from the uncrewed aircraft (i.e., be situated in a nearby crewed aircraft). Although the American radar-guided Bat glide bomb offered the possibility of fully autonomous operation, the onboard active radar homing terminal guidance system was not capable of reliably engaging the intended target and was susceptible to being thrown off by surface clutter. As a result, this radar-guided bomb was heavily reliant on the launch aircraft's larger and more capable human-operated radar, and the onboard radar guidance system was practically limited to naval applications wherein the sea surface constituted a less complex background environment than that encountered when attacking targets on land.

[21] Alastair Anthony, Norman K. Walker, and Elizabeth DeSocio, *The Accuracy of the Azon Guided Bomb as Affected by Battle Conditions in World War II* (Bethesda, MD: Norman K. Walker Associates), https://apps.dtic.mil/sti/pdfs/AD0600601.pdf; Donald I. Blackwelder, "The Long Road to Desert Storm and Beyond: The Development of Precision Guided Bombs," (Master's Thesis, Advanced Airpower Studies, Air University, 1992); Carlo Kopp, "The Dawn of the Smart Bomb, Air Power Australia," March 26, 2011, https://www.ausairpower.net/WW2-PGMs.html

To develop a more comprehensive appreciation of the implications of these advances in uncrewed aircraft technology, it is important to not limit ourselves to a monadic analysis. As previously alluded to, there was scope for the development of countermeasures to these early guided munitions. For example, although the German Fritz X enjoyed notable successes at the time of its initial combat employment in 1943, sinking American and British warships as well as (defecting) Italian warships, the radio links these guided bombs relied upon were inherently susceptible to electronic countermeasures. Within several months, American and British ships actively jammed the radio command link that allowed the human operator onboard the launch aircraft to control the Fritz X in flight.[22] Lacking the combination of onboard automated navigation and flight control or an onboard sensor suited for terminal guidance, these ostensibly guided bombs missed their targets with increasing regularity such that the otherwise pioneering Fritz X was quickly reduced to irrelevance, a dynamic both expedited and exacerbated by Germany's growing inability to compete with the United States and the United Kingdom in the air (the Fritz X was launched by crewed bomber aircraft).[23] Ultimately, the expense and complexity of developing and manufacturing guided munitions meant that these uncrewed aircraft constituted an insignificant fraction of the aerial munitions expended during the war and the extent of technological immaturity was such that the systems employed in combat often missed their intended targets.

To avoid the inherent limitations of control systems reliant upon real-time radio transmission and real-time control by a remote human operator, militaries required fully

[22] Samuel J. Cox, "H-Gram 021: Operation Avalanche, Fritz X, and the Battle of Durazzo," Naval History and Heritage Command, September 18, 2018, https://www.history.navy.mil/content/dam/nhhc/about-us/leadership/hgram_pdfs/H-Gram_021.pdf

[23] German responses to electronic countermeasures were made irrelevant by the inability of German aircraft to operate in the vicinity of their intended targets without being shot down.

autonomous uncrewed aircraft designs. During the Second World War, Germany developed and deployed two such systems, the V-1 cruise missile and the V-2 ballistic missile. These operated without a human in the loop and, in very different ways, incorporated technological advances that greatly expanded the realm of possibilities for uncrewed aircraft technology that would shape the course of future technological development.

The V-1 cruise missile is primarily relevant for the present purposes not as a result of its limited level of onboard automation, which was not without precedent and left much to be desired,[24] but as a result of its incorporation of a pulsejet engine, a crude and inefficient early form of jet engine. This offered the V-1 cruise missile a high speed – relative to propeller-driven aircraft – that made it militarily useful despite its many limitations, including poor accuracy, as high speed allowed it to evade air defenses (late war advances in aircraft, radar, and anti-aircraft artillery technology, however, greatly diminished the advantages of higher speed nominally conferred through the use of a pulsejet engine). Stated differently, the V-1 cruise missile serves as an example of an uncrewed aircraft design being appealing to militaries not as a result of an advance in automation technology but by incorporating a new component technology, in this case, a crude pulsejet engine, that could not practically be employed on necessarily larger and heavier crewed aircraft (in time, however, more efficient and higher power density turbojets facilitated the introduction of faster subsonic crewed fixed-wing aircraft).[25] A propeller-driven counterpart to the V-1, using much the same crude autopilot system,

[24] Stephen G. Evans and Keith B. Delaney, "The V1 (Flying Bomb) Attack on London (1944-1945); the Applied Geography of Early Cruise Missile Accuracy," *Applied Geography* 99 (2018): 44-53.
[25] Sterling Michael Pavelec, *The Jet Race and the Second World War* (Westport, CT: Prager, 2007)

could have been developed but – leaving aside the issue of (in)accuracy – would not have been as practical or effective as it would have been just as vulnerable to air defenses and enemy aircraft as any propeller-driven crewed aircraft while lacking the bounded capacity of an onboard human pilot to navigate and direct munitions to the intended targets in response to rapidly changing circumstances.

The V-2 ballistic missile is primarily relevant for the present purposes not as a result of its incorporation of breakthroughs in liquid rocket motor technology but because it incorporated an inertial navigation system and, unlike the V-1, thereby constitutes a specimen of uncrewed aircraft technology equipped with an onboard guidance system that would in later, more refined forms, allow for a practical means of operation beyond remote control over radio. To accurately strike a target while moving, all aircraft, whether crewed or uncrewed, and indeed all vehicles, need to know their present location relative to their starting location and their intended location. In crewed aircraft, pilots make such determinations with the aid of visual references, compasses, maps, watches to measure time, and onboard gauges. Whereas the V-1 was, aside from its pulsejet engine, a fairly rudimentary design employing a fairly crude autopilot system, the V-2 ballistic missile incorporated what was for its time a very sophisticated guidance system using gyroscopes, an accelerometer, and an analog computer to adjust its trajectory while in powered flight.[26] Although the inertial navigation system employed on the V-2 was barely adequate for an inaccurate ballistic missile directed against targets some 300 kilometers away, it nevertheless heralded a future in which uncrewed aircraft could

[26] Michael Neufeld, *The Rocket and the Reich: Peenemunde and the Coming of the Ballistic Missile Era* (Cambridge, MA: Harvard University Press, 1996); T.D. Dungan, *V-2: A Combat History of the First Ballistic Missile* (Yardley, PA: Westholme Publishing, 2019)

benefit from an onboard navigation system that facilitated independent operation including the undertaking course corrections without a human in the loop.

12.4 Assessing the Implications of Technological Change, ca. 1900-1945

In Part Two of this dissertation, I explained that human-powered military submarines were *possible* as early as ca. 1800 and that self-propelled military submarines were *possible* by the mid-1860s. Nevertheless, the effects of advances in submarine technology across nearly the entirety of the nineteenth century were negligible if not non-existent as a result of the immaturity of submarine technology and, no less importantly, the availability of more practical and more effective substitutes for the roles that submarines were intended. We encounter a similar dynamic when examining the implications of advances in uncrewed aircraft technology in the ca. 1910 through 1945 timeframe. By the end of the Second World War, uncrewed aircraft were not only distinctly *possible*, but militaries developed and employed uncrewed aircraft in a variety of roles in combat conditions. These uncrewed aircraft were, for the most part, unsuited to substituting for, let alone supplanting, crewed military aircraft but nevertheless allowed militaries to undertake certain new roles (for example, long-range uncrewed aerial bombardment) and, in favorable (i.e., permissive) conditions, undertake certain existing roles (i.e., crewed aerial bombardment) with greater effectiveness. For these reasons, we find very limited evidence that the development of increasingly practical uncrewed aircraft engendered sizeable and abrupt shifts in the fielded military capabilities of states in the ca. 1910 to 1945 timeframe.

Chapter 13: Uncrewed Aircraft Technology ca. 1945-1970

Part Three of this dissertation has thus far devoted considerable attention to advances in uncrewed aircraft technology made between ca. 1910 and 1945 that, all things considered, amounted to little in terms of shifts engendered in fielded military capabilities. The primary purpose of the analytical exercise in the previous chapter was to establish and cement a baseline through which to discern the pace and scale of the effects engendered by later technological developments. Without anchoring uncrewed aircraft technology in the first half of the twentieth century, there is a risk of truncating the developmental history in a manner that excludes decades of protracted and largely ineffectual technological developments such that the later emergence of more practical large reusable fixed-wing uncrewed aircraft designs may be inaccurately interpreted as abruptly engendering significant effects on the fielded military capabilities of states. While advances in uncrewed aircraft technology had negligible effects on the fielded military capabilities of states through 1945, various technological developments nevertheless expanded the technological frontier. The effects of technological change tend to be lagged and experiences in the ca. 1945-1970 timeframe may offer evidence in support of the proposition that technological change is a suitable candidate for a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

A comprehensive examination of the many manifestations of uncrewed aircraft technology developed in the years ca. 1945 through 1970 is beyond the scope of this study but it is productive to first examine several areas of technological development briefly in order to establish a baseline through which to analyze the effects of advances in large reusable fixed-wing (LRFW) uncrewed aircraft technology in this timeframe. In

addition to illustrating the realm of possibility concerning uncrewed aircraft technology, the survey of technological developments offered in the following pages will provide evidence that leading militaries were, by and large, eager to adopt uncrewed aircraft technology, broadly defined, when and where it was practicable. As a result, we can disabuse ourselves of any notion that leading militaries in the ca. 1945 to 1970 timeframe were uninterested in or opposed to uncrewed aircraft capable of undertaking the roles for which existing and prospective types of crewed combat aircraft were employed. Unless we can find evidence of *practical* LRFW uncrewed aircraft that could realistically compete with, and thereby substitute for, crewed aircraft in one or more roles, we will have no evidence that militaries in the ca. 1945 to 1970 timeframe *could have* but *chose not* to supplant crewed combat aircraft with uncrewed aircraft technology. As such, technological immaturity would constitute the primary explanation for the limited observable military effects – by virtue of non-deployment or deployment on a very small scale – of advances in uncrewed aircraft technology on the fielded military capabilities of states in the ca. 1945-1970 timeframe.

13.1 The Pursuit of Increasingly Automated Aircraft, Crewed and Uncrewed

13.1.1 The Eager Pursuit of Aircraft Automation: Complements to Substitutes for Nuclear-Armed Crewed Bomber Aircraft

The centrality of nuclear weapons to military planning in the early Cold War resulted in large and sustained investments toward the development of new and improved means of delivering nuclear weapons. Nuclear weapons are beyond the scope of this dissertation but developments in nuclear delivery systems are important for the present purposes

because the further development of two manifestations of uncrewed aircraft technology, ballistic and cruise missiles, resulted in advances in the areas of aircraft automation and guidance technology that would, in time, expand the realm of possibility for LRFW uncrewed aircraft technology to increasingly challenge crewed combat aircraft for funding and roles.

Although ballistic missiles would, in time, become the primary and most important means of nuclear delivery, this dynamic was not cemented until the 1960s. Until functional long-range ballistic missiles – with so-called intercontinental range given the geography of the American-Soviet nuclear competition – could be developed and with demonstrable reliability (as determined through extensive testing), the United States and the Soviet Union were willing to allocate resources toward the development of long-range cruise missiles to both complement and substitute for crewed bomber aircraft in delivering nuclear weapons.[27]

To understand the implications of advances in cruise missile technology for the practicality of LRFW uncrewed aircraft in the period ca. 1945 through 1970, we need to briefly return to the cruise missiles developed during the Second World War. If defined as a guided projectile powered in sustained flight by a reaction engine, then the German V-1 was the first cruise missile to be developed and deployed. Although the V-1 was groundbreaking, the United States was able to quickly reverse engineer the V-1 in the form of the Republic-Ford JB-2 and continued to develop this design following the end of

[27] William S. Borgiasz, *The Strategic Air Command: Evolution and Consolidation of Nuclear Forces, 1945-1955* (Westport, CT: Praeger, 1996); John M. Curatola, *Bigger Bombs for a Brighter Tomorrow: The Strategic Air Command and American War Plans at the Dawn of the Atomic Age, 1945-1950* (Jefferson, NC: McFarland & Company, 2016); Edward Kaplan, *To Kill Nations: American Strategy in the Air-Atomic Age and the Rise of Mutually Assured Destruction* (Ithaca, NY: Cornell University Press, 2015).

the war.[28] While the JB-2 refined the technology of the V-1, perhaps the greatest improvement made to the V-1 by American designers was realized after the war in the form of a new guidance system that relied on ground-based radars to track the missile in flight and provide course corrections over radio.

This new guidance system lay at the intersection of several developments in the navigation and guidance technologies that were employed by crewed bomber aircraft during the Second World War. As noted in the previous chapter, aerial bombing, particularly bombing undertaken at night and/or at high altitudes, tended to be highly inaccurate during the Second World War. Leaving aside the development and employment of guided bombs, the primary approach to addressing this challenge was to rely upon radio navigation systems to guide crewed bomber aircraft toward the intended target area. While Germany was the pioneer in this approach to increasing the accuracy of aerial bombing, the British military responded with an intense measure-countermeasure competition, the so-called Battle of the Beams, that degraded German radio navigation systems. The United Kingdom also developed and deployed radio navigation systems of its own including Gee, the first operational hyperbolic radio navigation system. Gee, and other systems that followed by the end of the war, namely the British Decca and American LORAN systems, offered aircrews what was essentially an offboard navigation system that allowed crewed aircraft to undertake more accurate bombing missions including nighttime missions for which no visual references were available to assist in navigation.

[28] "Republic/Ford JB-2 Loon (V-1 Buzz Bomb)," National Museum of the United States Air Force, No Date, https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/196227/republicford-jb-2-loon-v-1-buzz-bomb/

Toward the end of the war, the capabilities offered by radio navigation systems and the ubiquity of radar in American service led to experiments with the use of ground-based radar to track crewed bomber aircraft and JB-2 cruise missiles in flight and issue course corrections to the missile over radio. By so doing, designers developed what amounted to another offboard guidance system, one that had the potential for both increasing the accuracy of bombing conducted by crewed combat aircraft – ground-directed bombing – and offering a means for more accurate bombing by uncrewed aircraft in the form of cruise missiles.

Ground-directed bombing was extensively utilized by American crewed combat aircraft from the Korean War onwards and by 1952,[29] the United States deployed the first cruise missile designed to rely upon an external ground-based guidance system, the MGM-1 Matador.[30] Powered by a turbojet engine and equipped with an autopilot, the initial versions of the Matador relied on a derivative of the LORAN ground-based radio navigation system, but later versions relied on ground-based radars to undertake ground-directed bombing in much the same manner as then-contemporary crewed American bomber aircraft. Although the German V-2 ballistic missile had pioneered the use of inertial navigation systems (INS) in uncrewed aircraft, the INS systems available in the late 1940s and early 1950s lacked the accuracy required to deliver a relatively small nuclear warhead to a distance of several hundred kilometers on a slower cruise missile.[31]

---

[29] John R. Bruning, *Crimson Sky: The Air Battle for Korea* (London, UK: Brassey's, 1999)

[30] "Martin SSM-A-1/B-61/TM-61/MGM-1 Matador," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-1.html ; "Matador," Designation Systems, No Date, https://nuke.fas.org/guide/usa/theater/matador.htm

[31] All inertial navigation systems are subject to drift – increasing inaccuracy over time and distance.

The post-war application of ground-directed bombing to uncrewed aircraft allowed for the development of increasingly automated uncrewed aircraft such as the Matador that relied upon offboard systems for navigation and guidance. While it ostensibly addressed the problem of cruise missile inaccuracy, an issue exemplified by the decidedly mixed wartime record and effects of the otherwise innovative V-1, ground-directed bombing was inherently vulnerable to electronic interception and jamming, and a cruise missile – and any other type of aircraft – had to remain within the line-of-sight range of the relevant radars and radio stations for this means of off-board navigation and guidance to function. To address these shortcomings, the United States developed a successor to the MGM-1, the MGM-13 Mace, a nuclear-armed cruise missile that entered service in 1959.[32] This cruise missile is relevant for the present purposes because it incorporated both an inertial navigation system and an onboard ground-mapping radar.

As part of their efforts to develop longer-range ballistic missiles and crewed spaceflight, the United States and the Soviet Union undertook major efforts to advance inertial navigation technology such that, in time, smaller, lighter, and more accurate systems could be used on a much smaller cruise missile like the Mace.[33] Whereas the application of ground-directed bombing to a cruise missile forced the uncrewed aircraft to remain within line-of-sight of the relevant ground-based radars and radio stations, an inertial navigation system offered an uncrewed aircraft the ability to independently fly

[32] "Martin TM-76/MGM-13/CGM-13 Mace," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-13.html ; "Mace," Federation of American Scientists, No Date, https://nuke.fas.org/guide/usa/theater/mace.htm

[33] Donald MacKenzie, *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance* (Cambridge, MA: MIT Press, 1990); W. Wrigley, "History of Inertial Navigation," *Navigation* 24, No. 1 (1977): 1-6.; Thomas Wildenberg, "Origins of Inertial Navigation," *Air Power History* 63, No. 4 (2016): 17-24.

toward – and possibly return from – a target area and fly to preprogrammed waypoints along the way.

Just as ground-directed bombing and inertial navigation systems originated in the Second World War, the MGM-13's introduction of an onboard ground-mapping radar was an adaptation of the latest versions of a Second World War area of technological development to uncrewed aircraft. In 1943, the United Kingdom deployed the H2S, the first airborne ground-mapping radar, as a navigation aid for crewed bombers undertaking otherwise "blind" nighttime bombing missions.[34] While the British military used radio navigation systems including Gee, reliance on a ground-based radio navigation system restricted the practical range of bomber operations. The development of the cavity magnetron and higher frequency centimetric radar provided the British military with smaller airborne radars that were effective in detecting other aircraft. Experimentation with directing such a radar toward the ground led to the development of ground-mapping radars that allowed bombers operating at night to locate large targets such as cities, thereby overcoming the range restrictions imposed by the Gee ground-based radio-navigation system. The H2S radar, which operated in the S-band, was followed by the American H2X, a comparable but higher resolution radar operating in the higher frequency X-band. The ground-mapping radar used on the MGM-13 Mace, the Automatic Terrain Recognition and Navigation (ATRAN) system, built upon the foundations laid by these Second World designs and was the first to be adapted for use on an uncrewed aircraft in the form of a cruise missile (in later cruise missile designs,

[34] Bernard Lowell, *Echoes of War: The Story of H2S Radar* (Boca Raton, FL: CRC Press, 1991)

evolved forms of this approach to onboard aircraft navigation generally become known as Terrain Contour Matching, or TERCOM).

13.1.2 The Eager Pursuit of Aircraft Automation: Crewed Interceptor Aircraft and Ground-Based Air Defenses

The advent of jet propulsion in the form of subsonic, transonic, and later supersonic crewed aircraft increased the challenge of air defense and undermined the effectiveness of many areas of wartime developments in air defenses. Given that the nuclear threat was essentially synonymous with nuclear-armed bomber aircraft for most of the first two decades of the Cold War and that their nuclear arsenals were primarily aimed at one another, the United States and the Soviet Union converged upon two applications of aircraft automation technology for the role of air defense: the development of highly automated crewed interceptor aircraft that relied upon ground-based radars for what amounted to off-board fire control to vector interceptors toward enemy bomber aircraft, and the development of long-range guided surface-to-air missiles to both complement and substitute for crewed fighter and interceptor aircraft in air defense roles over large urban areas and key industrial and military facilities that were priority targets for nuclear-armed bomber aircraft in a prospective war.

In the early Cold War, leading militaries, including both the United States and the Soviet Union, built upon pioneering wartime research into rocket-powered guided surface-to-air missiles. In 1954, the United States deployed the MIM-3 Nike Ajax, the first operational surface-to-air missile system. The Soviet Union, in turn, deployed its first operational system, the S-25, in 1955 (i.e., the SA-1, to use its NATO designation).

These systems constituted the practical application of rocket propulsion technology to air defense after the less-than-stellar employment of the rocket-powered German Me 163 crewed short-range interceptor aircraft during the Second World War.

Rocket propulsion offered excellent acceleration and the potential for a high maximum speed that even supersonic crewed aircraft (using air-breathing jet engines) struggled to attain, but rockets were impractical in crewed aircraft as the burn time of rockets was measured in seconds and a rocket-powered aircraft would afterward coast. The pilots of the Second World War German Me 163 interceptor would therefore climb to altitude while the rocket was burning and then trade altitude for speed to engage bomber aircraft before having to land what amounted to an unpowered glider so that it could be refueled for another mission. Without a human pilot onboard, a rocket-propelled aircraft could be employed as a single-use system that crashed into a target aircraft (or detonated nearby if appropriately fused) and could additionally escape the design restrictions imposed by human physiological limitations. As a result, an uncrewed rocket-propelled single-use aircraft used in an anti-aircraft role could operate at a speed – and be subject to forces of acceleration – that far exceeded what any human pilot could sustain and, with suitable control surfaces, could undertake maneuvers that similarly exceeded what any human pilot could manage.

As with other types of uncrewed aircraft, the primary challenge to developing a guided anti-aircraft missile was finding a guidance system that could effectively substitute for an onboard human pilot. The advent of higher frequency radar offered a (ground-based) means of detecting and tracking aircraft and, when coupled with a means of computation to determine a fire control solution, a radio link, and a second (ground-

based) radar to track the missile in flight, the missile could be guided toward the target aircraft and detonated through command instructions transmitted over radio (i.e., early anti-aircraft missile would be guided from the ground and did not have an onboard radar antenna to receive let alone transit radar emissions). By deploying such surface-to-air missile systems, militaries increasingly transitioned away from anti-aircraft artillery and short-range interceptor aircraft for air defense over urban areas and key industrial and military facilities.

While surface-to-air missile systems had considerable potential, these systems were not ready for military service until the mid to late 1950s, and early systems had major limitations and could not be solely relied upon to provide effective air defense over large countries such as the United States and the Soviet Union. Early missiles were also primarily effective against larger and less maneuverable aircraft such as bombers and could not target aircraft more than several dozen kilometers away. The line-of-sight nature of essentially every radar in existence and the fact that early surface-to-air missiles were command-guided systems (i.e., not fire-and-forget systems with onboard guidance systems) meant that the practical engagement range was largely determined by the altitude of the target aircraft such that bombers could evade radar coverage – and surface-to-air missile systems – by flying at lower altitudes. Given the severity of the threat posed by nuclear-armed crewed bomber aircraft, the United States and the Soviet Union developed and deployed larger, heavier, and increasingly automated crewed interceptor aircraft to engage enemy bombers as far away as possible.

Wartime advances in radar technology had allowed militaries to equip crewed combat aircraft with onboard radars capable of detecting airborne targets. While initially

limited to large night fighters, the advent of higher frequency centimetric radar enabled by the development of the cavity magnetron allowed for the development of radar-equipped fighter and interceptor aircraft. The size and weight of early radars, determined as it was by the reliance on electronics using vacuum tubes, created a situation in which only larger, heavier, and generally slower and less maneuverable crewed combat aircraft could practically employ air-to-air radar. After the war, militaries prioritized increasing aircraft speed and maneuverability through the adoption of turbojet engines, but early turbojets left much to be desired in terms of power density and fuel efficiency. As a result, early jet-powered fighter aircraft were not equipped with air-to-air radar and militaries instead developed larger, heavier crewed interceptor aircraft that were optimized for high speed – the characteristics associated with guided surface-to-air missiles – at the expense of maneuverability and range in order to engage bomber aircraft on their expected avenues of approach.

While crewed interceptor aircraft were equipped with radar, these radars had a limited field of view and limited tracking capabilities. Air defense operators on the ground, who monitored vast volumes of airspace and increasingly benefitted from the networking of disparate radars, had far greater awareness of aerial activity notwithstanding the constraining effects of the radar horizon on line-of-sight radar systems. As a result, early Cold War interceptor aircraft were, much like their Second World War counterparts, primarily reliant on ground-controlled interception which is to say that a crewed interceptor aircraft would take off, follow vectors communicated via radio by ground-based air defense operators who oversaw the ground-based radar network, and, upon direction of the ground controllers, would activate an onboard radar

(if present) and attack detected bomber aircraft. In the context of advances in automation technology, leading militaries, particularly the United States military, worked to automate the ground-controlled interception process such that the crewed interceptor aircraft could automatically follow vectors transmitted over radio so that the pilot was largely reduced to the role of a weapons system operator during the interception mission. By the late 1950s, the United States deployed the crewed F-106 Delta Dart interceptor alongside the Semi-Automatic Ground Environment (SAGE) integrated air defense system which together constituted a level of automation that essentially turned crewed interceptor aircraft into a form of reusable surface-to-air missile used to launch guided air-to-air missiles against crewed bomber aircraft before returning to base for rearming and refueling.

13.1.3 The Eager Pursuit of Aircraft Automation: Fighter Aircraft

Although fighter aircraft were at the forefront of many areas of advancement in military aviation in the early Cold War period, the computers used to automate the operations of interceptor aircraft were too large and heavy for use on smaller and lighter fighter aircraft throughout much of this timeframe and were therefore limited to larger crewed bomber and interceptor aircraft. While this dynamic would later change such that fighter aircraft designs would also benefit from advances in automation technology from the mid-1960s onward, by the late 1950s, militaries began to field the latest air-to-air missiles on radar-equipped fighter aircraft and in so doing further demonstrated their eagerness to adopt the latest advances in an area of uncrewed aircraft technology – guided munitions – when and where possible. Such was the confidence of the United States military, which plays a

central role in the development of large fixed-wing uncrewed aircraft technology in this timeframe, in its air-to-air missile technology that it introduced a new crewed aircraft, the F-4 Phantom twin-engine heavy fighter, without an internal gun for use in close range aerial engagements.[35] The United States military did so based on its erroneous assessment that the combination of the latest airborne radar and the infrared-guided AIM-4 Falcon, the first operational (guided) air-to-air missile, and the longer-range semi-active radar-homing AIM-7 Sparrow, had made aircraft guns and close-range air combat maneuvering (i.e., "dogfighting") less relevant.[36] Despite extensive testing in the decade or so leading up to 1965, both guided air-to-air missiles (as well as the related AIM-9 Sidewinder developed by the United States Navy) proved to be unreliable and ineffective when used against Soviet combat aircraft operated by North Vietnamese pilots, a dynamic that, among other things, resulted in the U.S. military retraining its pilots in air combat maneuvering and the development and manufacture of new versions of the F-4 fighter with a built-in gun for close-range aerial engagements.[37]

13.1.4 The Eager Pursuit of Automation: Guided Anti-Ship Missiles

Enthusiasm for and confidence in the effectiveness of the latest advances in uncrewed aircraft technology also drove the development and adoption of anti-ship missiles. Whatever its efficacy and morality, two matters both stemming from the guaranteed loss

[35] Richard P. Hallion, "A Troubling Past: Air Force Fighter Acquisition Since 1945," *Airpower Journal* 4, No. 4 (1990): 4-23.
[36] Sean O'Connor, "Arming America's Interceptors: The Hughes Falcon Missile Family," Air Power Australia, June 2011, https://www.ausairpower.net/Falcon-Evolution.html
[37] Marshall L. Michel III, *Clashes: Air Combat over North Vietnam 1965–1972* (Annapolis, MD: Naval Institute Press, 1997); Craig C. Hannah, *Striving for Air Superiority: The Tactical Air Command in Vietnam* (College Station, TX.: Texas A&M University Press, 2002)

of a trained human pilot, Imperial Japan's use of kamikaze aircraft against American surface ships indirectly highlighted the potential of uncrewed aircraft technology if only militaries could develop automated systems that could fully substitute for the onboard pilot in terms of navigating the aircraft, identifying a potential target, and thereafter guiding the aircraft laden with explosives into the hull of a surface ship. As the previous pages indicate, the period ca. 1945-1970 timeframe was characterized by the widespread development and deployment of increasingly automated aircraft, whether crewed or uncrewed. Given this backdrop, militaries unsurprisingly pursued the development of uncrewed fixed-wing aircraft that could wholly supplant the capabilities offered by crewed kamikaze aircraft in the anti-ship role and, in so doing, reduce the exposure of crewed aircraft in the face of advances in naval air defense capabilities. While the United States had been at the forefront of the development of guided air-to-surface munitions capable of targeting surface ships during the Second World War, it did not face a remotely comparable surface fleet after 1945 and it was the Soviet Union, a country that faced the qualitatively and quantitively superior American fleet, that led the way in developing and deploying guided anti-ship missiles.

From the mid-1950s onward, the Soviet Union deployed a growing variety of aircraft-, ground-, ship-, and submarine-launched guided anti-ship missiles of which the air-launched KS-1 and surface-launched P-15 are broadly representative early Cold War examples. Derived from technology developed for the crewed MiG-15 fighter, the earlier of these two designs, the turbojet-powered KS-1, which weighed around 3,000 kilograms,[38] was similar in appearance to a small fixed-wing crewed jet-powered aircraft.

[38] This was around half the maximum take-off weight of a crewed MiG-15 fighter.

Launched from the wings of a crewed Tu-16 medium bomber, the KS-1's autopilot navigated the uncrewed aircraft toward the target ship until the target ship was illuminated by the Tu-16's radar (i.e., there was a human in the loop), after which the KS-1's onboard semi-active radar homing seeker provided terminal guidance and directed the missile to the desired area of impact. Launched from altitude, the KS-1 allowed a crewed Tu-16 bomber to engage a surface warship at distances of fifty or more kilometers such that the Tu-16 would be immune to anti-aircraft gunfire and largely beyond the effective range of early shipborne surface-to-air missiles, a dynamic that led the United States to pursue new fleet air defense fighters that culminated in the introduction of the F-14 fighter in the early 1970s.

Whereas the KS-1 was an early and rather primitive air-launched anti-ship missile, the Soviet P-15 (NATO designation: Styx – a later better-known Chinese version is designated the Silkworm), a 2,500-kilogram rocket-powered design that entered service in 1960, was a more refined surface-launched anti-ship missile that is reflective of advances in missile guidance made throughout the 1950s. The P-15 did not benefit from being launched from altitude and at relatively high speed from a jet-powered bomber aircraft in the manner of the KS-1 and initial versions offered a maximum range of just 40 kilometers (the P-15 did not have an air-breathing jet engine and much of its considerable weight and volume consisted of fuel and oxidizer for a liquid fuel rocket engine). In this respect, the P-15 may appear to have been an inferior anti-ship missile but is particularly it relevant for the present purposes because it was equipped with a terminal active radar homing guidance system to complement the (more refined) autopilot and an onboard inertial navigation system. Unlike the KS-1, which had no onboard means of

tracking and hitting a target ship without said ship being illuminated by the launch aircraft's radar, the P-15's onboard active radar homing guidance system allowed for fully independent and automated operation in the same manner as the wartime American ASM-N-2 Bat.

Early Soviet anti-ship missiles such as the KS-1 had major limitations but successive technological advances throughout the 1950s and 1960s meant that anti-ship missiles, whatever their launch platform, constituted an increasingly credible threat to surface warships by the late 1960s, a dynamic brought to international attention by the sinking in 1967 of a (Second World War vintage) Israeli destroyer by a Soviet P-15 anti-ship missile launched from an Egyptian missile boat. The net effect of these gradually unfolding technological advances is that advances in aircraft automation and guided missile technology allowed militaries to deploy large non-reusable fixed-wing uncrewed aircraft, whether launched from an aircraft, surface ship, submarine, or from the ground, in a manner that supplanted crewed fixed-wing aircraft in the role of (pilotless) kamikaze aircraft and aerial anti-shipping operations more generally.

What these brief examples highlight is that the period ca. 1945-1970 was one in which militaries were eager to adopt the latest in uncrewed aircraft technology which, being broadly conceptualized for the purposes of this dissertation, encompasses both the automation of crewed aircraft and the employment of guided munitions from crewed aircraft and other launch platforms. Given the extensive evidence for this dynamic across multiple countries, including both the United States and the Soviet Union, there is little reason to think that the militaries of this timeframe would have been unreceptive to proposals for large reusable fixed-wing uncrewed aircraft designs capable of substituting

for if not supplanting crewed aircraft in one or more roles. At the same time, of course, successive advances in crewed aircraft technology and other areas of uncrewed aircraft technology meant that proposals for large reusable fixed-wing (LRFW) uncrewed aircraft encountered increasing competition for funding and roles.

13.2 The Development of Large Reusable Fixed Wing Uncrewed Aircraft in the United States

While leading militaries eagerly adopted increasingly automated crewed combat aircraft and an increasingly diverse array of uncrewed aircraft in the form of air-to-surface, air-to-air, surface-to-air, and surface-to-surface guided munitions, these technological developments were primarily aimed at augmenting human pilots and thereby complementing, not substituting for, crewed military aircraft. By the late 1950s, however, the United States identified a potential application for LRFW uncrewed aircraft technology in which an uncrewed aircraft would substitute for a crewed aircraft in an important role. Understanding the context in which these developmental objectives were pursued is important because the late 1950s marked the beginning of what would later become a close association between LRFW uncrewed aircraft technology and intelligence, surveillance, and reconnaissance (ISR) roles.

American concerns about Soviet nuclear capabilities led to an expansive multifaceted intelligence-gathering program including a far-reaching aerial reconnaissance effort.[39] Intelligence-gathering flights along the Soviet Union's borders

[39] Krzysztof Dabrowski, *Hunt for the U-2: Interceptions of Lockheed U-2 Reconnaissance Aircraft over the USSR, Cuba and People's Republic of China, 1959-1968* (Warwick, UK: Helion and Company, 2020);

and coastline typically made use of large crewed aircraft, such as the reconnaissance versions of bomber, transport, and maritime patrol aircraft, that were suitable for longer-range and/or higher endurance missions.[40] The large size of the electronic devices (which used vacuum tubes) required for many ISR missions necessitated the employment of large aircraft and the large enclosed volume and payload of such aircraft was desirable as it allowed a group of intelligence analysts to operate onboard the aircraft and begin processing intelligence in real-time, thereby allowing for the redirection of intelligence gathering efforts in response to new information. In the absence of a practical means for transmitting large volumes of data over long distances given the limitations of radio (datalink) technology, the most capable ISR aircraft of the day required pilots, analysts, and technicians onboard to undertake the roles that could not be automated or performed remotely.

The immense expanse of the Soviet Union, the sheer number of potential avenues of approach, and the not-unimportant fact that the Soviet Union had been a laggard in the development of radar-centric air defense capabilities during the Second World War, all contributed to a situation in which the United States could employ long-range (bomber) aircraft configured for ISR missions on overflights inside much of Soviet airspace with near total impunity.[41] By the mid-1950s, however, this window of opportunity was

---

Gregory Pedlow, Donald Welzenbach, and Chris Pockock, The Central Intelligence Agency and Overhead Reconnaissance: The U-2 and OXCART Programs, 1954-1974 (New York: Skyhorse, 2016); Dino A. Brugioni, *Eyes in the Sky: Eisenhower, the CIA and Cold War Aerial Espionage* (Annapolis, MD: Naval Institute Press, 2010)

[40] John T. Farquhar, *A Need to Know: The Role of Air Reconnaissance in War Planning, 1945-1953* (Maxwell Air Force Base, AL: Air University Press)

[41] Walter J. Boyne, "The Early Overflights," Air & Space Forces Magazine, June 1, 2001, https://www.airandspaceforces.com/article/0601overfly/ ; Kevin Wright, *We Were Never There: Volume 1: CIA U-2 Operations Over Europe, the USSR, and the Middle East, 1956-1960* (Warwick, UK: Helion and

rapidly closing on the heels of the Soviet Union's deployment of ever-increasing numbers of air defense radars, jet-powered fighters and interceptor aircraft, and, from 1957 onward, long-range guided anti-aircraft missiles. In this context, the United States developed the U-2, an intelligence-gathering aircraft with glider-like high aspect ratio wings optimized for operations at much higher altitudes that would, in principle, allow the aircraft to conduct overflights and remain beyond the reach of Soviet fighter and interceptor aircraft.

While progressive advances in Soviet air defense capabilities meant that U-2 overflights were increasingly detected by Soviet radars, the Soviet Union was unable to successfully intercept any U-2 aircraft until the high-profile shooting down of a U-2 in 1960.[42] The diplomatic fallout that ensued, the exposure of the U-2's technology, and the capture of the downed U-2's pilot led to the termination of American (crewed) ISR missions inside Soviet airspace. Although this decision hindered American intelligence gathering efforts for a time, the repercussions were short-lived and were quickly counteracted by the success of the pioneering American CORONA imaging satellite program, which provided relatively high-resolution imagery of the Soviet Union from the early 1960s onwards.[43]

With the advent of uncrewed ISR satellites, a new – uncrewed – military technology seemingly supplanted existing crewed military aircraft in an important role.

Company, 2021).; Kevin Wright, *We Were Never There: Volume 2: CIA U-2 Asia and Worldwide Operations 1957-1974* (Warwick, UK: Helion and Company, 2021).

[42] Steven J. Zaloga, "Defending the Capitals: The First Generation of Soviet Strategic Air Defense Systems 1950–1960," *Journal of Slavic Military Studies* 10, No. 4 (1997): 30-43; Steven J. Zaloga, *Soviet Air Defence Missiles: Design, Development and Tactics* (London, UK: Jane's Information Group, 1989)

[43] Dwayne A. Day, John M. Longdon, and Brian Latell, editors. *Eye in the Sky: The Story of the Corona Spy Satellites* (Washington, D.C.: Smithsonian Institution Press, 1998)

For all the advantages the new technology offered relative to extant ISR aircraft, however, satellites were not suitable for all ISR roles, and the nature of orbital mechanics (in low earth orbit) meant that satellites offered a new niche ISR capability rather than a range of ISR capabilities that could replicate what crewed ISR aircraft offered militaries. Fixed-wing atmospheric aircraft remained effective in many peacetime ISR roles, and the United States remained particularly interested in ISR aircraft that could undertake missions on short notice when and where required, unlike a satellite with its inflexible (low earth) orbits and fleeting if regular, overhead transits that routinely captured unusable optical band imagery as a result of cloud cover. In addition to camera-equipped crewed fighter and bomber aircraft, which were primarily used for so-called tactical ISR roles, and the U-2, which continued to operate outside of Soviet airspace, the United States pursued four new ISR aircraft designs to complement the capabilities offered by the combination of existing crewed ISR aircraft and the new ISR satellites. Three of these new ISR aircraft are important for the present purposes because these did not have a human operator onboard and thereby constitute the first large fixed-wing uncrewed aircraft developed for ISR roles.

While the U-2 had many positive attributes, it lacked the range necessary to reach all target areas and return to an airfield. To compensate for this, the United States deployed U-2s, like other crewed ISR aircraft, from foreign bases that were closer to the Soviet Union. The U-2 shot down in 1960 over the Soviet Union, for example, took off from an airbase in Pakistan and was scheduled to land in Norway. Some areas of interest to the United States, above all nuclear facilities located in western China (i.e., Lop Nur), were beyond the reach of the U-2 and China's deployment of increasing numbers of

Soviet-designed S-75 surface-to-air missile systems (i.e., the SA-2) made U-2 operations increasingly dangerous over the 1960s (China shot down five U-2s operated by the nationalist government in Taiwan in this timeframe).[44] In this context, in 1966 the United States sought a large reusable fixed-wing uncrewed aircraft that could reach areas of interest in China and return so that the film containing the captured imagery could be recovered and do so without putting at risk any human operators.

13.2.1 The American AQM-34

The first uncrewed aircraft developed for this role was the AQM-34, also known as the Ryan Model 147 Lightening Bug.[45] Developed in the early 1960s on the basis of the Ryan Firebee, a turbojet-powered uncrewed target aircraft used for training and testing purposes, the AQM-34 was perhaps the first LRFW uncrewed aircraft developed for employment in an ISR role. Like the Firebee, the AQM-34 could be launched either from the ground with the assistance of a booster rocket in the manner of a cruise missile or from the wings of a version of the C-130 military cargo aircraft, the DC-130, and this second mode of employment increased both its range and operational flexibility. Equipped with a rudimentary autopilot in its earliest versions, the AQM-34 was designed to fly toward a target area, capture aerial imagery, and return toward a designated area where, after the deployment of a parachute, the aircraft could be recovered so that the onboard wet film could be retrieved and developed.

---

[44] Bob Bergin, "The Growth of China's Air Defenses: Responding to Covert Overflights, 1949-1974," *Studies in Intelligence* 57, No. 2 (2013): 19-28.
[45] "Teledyne Ryan Q-2/KDA/xQM-34/BGM-34 Firebee," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-34.html

While initially employed on overflights over China, the AQM-34 lacked the range necessary to reach areas of interest deep inside Chinese airspace, a dynamic that drove the development of the second American LRFW uncrewed ISR aircraft, the AQM-91, a related design that was never operationally deployed. Although the AQM-34 later played a notable but limited role as an ISR platform in the context of the Vietnam War,[46] the United States turned to a new crewed aircraft design to supplant the U-2.

13.2.3 The American A-12/SR-71 and D-21

While the U-2's very high operating altitude allowed it to evade crewed Soviet fighter and interceptor aircraft, the U-2 was not a particularly fast aircraft, and advances in Soviet surface-to-air missile systems, as well as later advances in air-to-air missiles carried by Soviet interceptor aircraft, meant that the crewed U-2 was increasingly unsuited to operating in increasingly contested airspace over countries that received military equipment from the Soviet Union. The uncrewed AQM-91, which was developed to substitute for the U-2 in high-risk missions, had the same limitations. In response, the United States developed a new high-altitude ISR aircraft, the Lockheed A-12, the predecessor to its better-known derivative, the SR-71, that could fly at even higher altitudes and a speed in excess of Mach 3.

While the development of the A-12 and SR-71, which, given their high operating speed, large turning radius, and inability to loiter, operated in a manner more akin to satellites than the U-2, was successful and offered the United States substantial gains in

[46] Steve Miller, *The 99th Strategic Reconnaissance Squadron: The Air Force's Story of Unmanned Reconnaissance in the Vietnam War* (Miller Management Systems, 2017); William Wanger, *Lightning Bugs and Other Reconnaissance Drones* (Fallbrook, CA: Aero Publishers, 1982)

terms of ISR capabilities, both were crewed aircraft. The loss of a U-2 over the Soviet Union and concerns about the trajectory of developments in Soviet air defense capabilities, which were increasingly also employed by other countries, led to the development of a smaller, but equally fast, uncrewed derivative of the technology used in the A-12 and SR-71. The resulting uncrewed aircraft, the Lockheed D-21, was intended to substitute for its crewed counterparts on increasingly dangerous missions in Chinese airspace.

When first flown in 1964, the D-21 constituted perhaps the most advanced fixed-wing uncrewed aircraft then in existence. Beyond its maximum speed and altitude, the D-21 incorporated an inertial navigation system derived from the guidance systems developed for nuclear-armed ballistic and cruise missiles.[47] This allowed the D-21 to fly on preprogrammed trajectories without a human in the loop and, not unimportantly, enabled the D-21, which was equipped with a wet film camera, to return from the target area toward friendly airspace such that the captured imagery could be recovered and developed for analysis. To obtain a maximum speed of over Mach 3, the D-21 was equipped with a powerful ramjet engine. Because ramjet engines require a supersonic airflow to properly function, the D-21 was originally designed to be launched from a pylon mounted on the top of a (supersonic) M-21 carrier aircraft, a version of the A-12. As a result, the D-21 had a very complex – and very expensive – launching procedure but air-launch from a (crewed) aircraft served to essentially extend the maximum range of the D-21 such that the D-21 was, like the AQM-34, smaller and lighter than an uncrewed

[47] "Lockheed D-21B," National Museum of the United States Air Force, No Date, https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/195778/lockheed-d-21b/

aircraft of similar performance otherwise would have been if it began flight at zero altitude and airspeed.[48]

Despite the great expense of its construction and operation, the D-21 was not a reusable aircraft as there were no practical means of recovering a supersonic aircraft in flight, and automation technology was not so advanced as to allow for automated landing on a runway. Instead, the D-21 was designed to return from its mission area, jettison its camera film (via parachute) over friendly airspace much like early ISR satellites, and thereafter self-destruct. Notwithstanding its glaring limitations, the D-21 was far removed from the radio-controlled uncrewed target aircraft that became commonplace in the 1950s and had more in common with contemporary guided missiles, above all cruise missiles, particularly in areas of navigation and guidance technology. Innovative though its design was, the D-21 constituted a complex and, according to publicly available information, impractical and ultimately unsuccessful advance in uncrewed aircraft technology that did not offer an appreciable return on the immense investment made toward its development, manufacture, and operation.

13.2.4 The Soviet Tu-123

While the United States developed a substantial lead in LRFW uncrewed aircraft technology during the ca. 1945-1970 timeframe, the Soviet Union, for its part, pursued a less ambitious high-altitude and high-speed uncrewed ISR aircraft of its own. The

[48] This launch procedure did not work as intended and D-21 uncrewed aircraft were typically launched from the underwing pylons of crewed subsonic B-52 bombers. A solid rocket motor was attached to the D-21 to bring the uncrewed aircraft to supersonic speed and thereby allow for the operation of the ramjet engine.

product of this undertaking, the Tu-123, was developed on the basis of Soviet efforts toward developing a (very) long-range supersonic nuclear-armed cruise missile before ballistic missile technology became the preferred means of substituting for crewed bomber aircraft in the delivery of nuclear warheads. As a result of this design ancestry, the ground-launched Tu-123 was an extremely large and heavy uncrewed aircraft weighing around 35,000 kilograms, around seven times as much as the (air-launched) American D-21.[49] Like the American D-21, the Tu-123 was a single-use system designed to fly toward an area of interest, capture aerial imagery, and then return toward the launch area so that an ejected capsule containing the wet film could be recovered. It was, therefore, an extremely expensive means of generating aerial imagery (much like the American D-21) and while it offered the Soviet Union an (atmospheric) high-altitude ISR capability that it did not previously possess, the uncrewed Tu-123 was shortly supplanted in frontline service by reconnaissance versions of crewed combat aircraft including reconnaissance versions of the high altitude supersonic MiG-25 interceptor, the aircraft from which the Tu-123's engine was derived, and was, like its American counterparts, subject to growing competition from imaging satellites for funding and roles.

13.3 Assessing the Implications of Technological Change, ca. 1945-1970

As the experience with the AQM-34 and other American LRFW uncrewed ISR aircraft demonstrates, it was *possible* to build uncrewed aircraft to take on some of the roles of crewed aircraft in the ca. 1945-1970 timeframe, but LRFW uncrewed aircraft were, by

[49] Jonathan Pike, "Tu-123," Global Security, No Date, https://www.globalsecurity.org/military/world/russia/tu-123.htm

and large, both impractical and inferior substitutes to crewed aircraft. Despite extensive and expensive research and development efforts in this timeframe, the enduring limitations of LRFW uncrewed aircraft technology meant that this category of uncrewed aircraft was most competitive for niche high-risk roles in which the loss of a human operator – *in peacetime* – was unacceptable for political and/or diplomatic reasons. Even the AQM-34, which was by far the most successful of the aforementioned designs, could not compete with crewed ISR aircraft for funding and roles and was exclusively employed in a comparatively small number of niche ISR roles. In a world of infinite resources, militaries may have made the development of such uncrewed aircraft a top priority and procured and fielded large numbers of existing LRFW uncrewed ISR aircraft like the AQM-34, but uncrewed ISR aircraft were expensive and limitations in automation technology made these broadly inferior substitutes for crewed ISR aircraft.

It is important to recognize that the absence of the large-scale adoption of LRFW uncrewed aircraft through 1970 does not indicate that militaries were uninterested in uncrewed aircraft technology. After all, leading militaries pursued many forms of uncrewed aircraft technology during the ca. 1945-1970 timeframe, from cruise missiles and surface-to-air missiles to anti-ship missiles, and both developed and deployed increasingly automated crewed aircraft. In this time limit, these were the practical applications of uncrewed aircraft technology, broadly defined, and proponents of LRFW uncrewed aircraft technology for ISR and other roles had to wait for substantial advances in aircraft automation, sensor, and radio communication technology that would not be realized until the 1970s. While there is limited evidence to be found in the ca. 1945-1970 timeframe that advances in uncrewed aircraft technology engendered sizeable and abrupt

shifts in the fielded military capabilities of states, it is possible that the military effects of technological developments made in this timeframe were considerably lagged such that the longitudinal analysis must continue in search of supporting evidence.

Chapter 14: Uncrewed Aircraft Technology ca. 1970-1990

While significant advances were made in LRFW uncrewed aircraft technology in the ca. 1945-1970 timeframe, the effects on fielded military capabilities were minor, LRFW uncrewed aircraft were not very practical or cost-effective and were, therefore, inferior substitutes to crewed aircraft and could not productively supplant the latter in most if not all roles. Nevertheless, the technological advances made in the ca. 1945-1970 timeframe expanded the technological frontier and the effects of technological change can be lagged. It is therefore productive to continue the longitudinal assessment of the military effects of advances in uncrewed aircraft technology to the ca. 1970-1990 timeframe. By the late 1960s, the United States was the undisputed leader in LRFW uncrewed aircraft technology and deployed several notable and innovative designs that reflected the realm of technological possibility by the turn of the decade. Nevertheless, three developments in the early 1970s led the United States to dramatically scale back its employment of and ambitions for uncrewed (atmospheric) aircraft in the ISR role, thereby all but reversing – eliminating – the effects of earlier advances in LRFW uncrewed aircraft technology on its fielded military capabilities.

14.1 Changing Requirements in the United States

The end of American involvement in the Vietnam War in 1973, which was preceded by the curtailing of first ground and later air operations, resulted in a situation in which the United States was no longer undertaking high-risk ISR missions over North Vietnam's defended airspace. With large numbers of more capable crewed ISR aircraft available and without the imperative to avoid the death or capture of a human pilot in peacetime,

demand for the services of the AQM-34 experienced a dramatic decline in the early 1970s such that this uncrewed aircraft's supporting role in Operation Linebacker II in late 1972 would be both perhaps its most impactful employment as well as its last major combat operation.[50] Although the development and trials of new versions of the AQM-34 attempted to make this uncrewed ISR aircraft more suitable for operations in other areas, above all in Central Europe, a combination of technological limitations, a less conducive military-geographical context, competition with existing and prospective crewed ISR aircraft (as well as uncrewed ISR satellites), and bureaucratic disputes[51] resulted in the AQM-34 being withdrawn from service by 1975.

While the end of American involvement in the Vietnam War affected demand for so-called tactical ISR missions, the normalization of American diplomatic relations with China in 1972 also eliminated the requirement for an uncrewed ISR aircraft capable of operating in Chinese airspace (i.e., undertaking so-called strategic reconnaissance missions). The United States had developed the AQM-91 and the D-21 for this specific undertaking after Chinese air defenses, equipped as they were with the Soviet S-75 surface-to-air missile system, succeeded in shooting down both uncrewed AQM-34 and crewed U-2 aircraft (the latter being operated by the nationalist government in Taiwan). Without further operations, crewed or uncrewed, in Chinese airspace, the (high-altitude) AQM-91 and (high speed and high altitude) D-21 directly competed with the crewed (high-altitude) U-2 and (high speed and high altitude) SR-71 for funding and roles. These uncrewed substitutes for crewed ISR aircraft were inferior in terms of every parameter but one – not having humans onboard to put at risk – and were as such uncompetitive

[50] Thomas P. Ehrhard, *Air Force UAVs: The Secret History* (Washington, D.C.: Mitchell Institute Press, 2010)
[51] Ehrhard, *Air Force UAVs.*

designs absent the imperative to avoid the death or capture of an onboard pilot, an imperative that essentially no longer existed in peacetime after the conclusion of United States' involvement in the Vietnam War and its normalization of relations with China. The trade-offs imposed by the limited state of aircraft automation technology were immense; the D-21 was a very expensive single-use cruise missile-like uncrewed aircraft and the AQM-91 was incapable of undertaking the complex peacetime and wartime operations of the crewed U-2, which benefitted from the navigation, guidance, and tactical decision-making of its onboard human pilot (the same could be said of the D-21 relative to the SR-71).

While demand for the types of ISR missions for which uncrewed aircraft were most suited – missions that called for zero risk of the death or capture of aircrew – was sharply reduced, concurrent advances in the American ISR satellite technology offered more capable and more reliable means – as well as an inherently safe means – of undertaking regular intelligence gathering activities over the Soviet Union, China, and elsewhere. In 1971, the United States launched its first KH-9 HEXAGON satellite, a replacement for the pioneering CORONA program satellites, and was conducting research and development efforts that would result in the 1976 launch of the first KH-11 KENNEN satellite, which was the first ISR satellite equipped with an electro-optical sensor (i.e., an ISR satellite that transmitted captured imagery over radio instead of ejecting capsules containing wet film that had to be recovered and then developed after re-entering the atmosphere).[52] The growing number of ISR satellites in orbit and the increasing sophistication of said satellites, two supply-side dynamics, offered the United

[52] Dwayne A. Day, "Intersections in Real Time: the Decision to Build the KH-11 KENNEN Reconnaissance Satellite (part 1)," The Space Review, September 9, 2019, https://www.thespacereview.com/article/3791/1

States the ability to readily undertake peacetime orbital "overflights" beyond the reach of air defenses without putting any human operators at risk.

Together, these three developments affected both demand-side and supply-side dynamics in the United States, the undisputed leader in large uncrewed aircraft technology, such that proposals for new LRFW uncrewed aircraft intended for ISR roles had to directly compete not only with the crewed U-2 and SR-71, among other crewed ISR aircraft in American service but also increasingly capable ISR satellites.

14.2 The Soviet Union's Development of LRFW Uncrewed Aircraft Technology

While the United States was abandoning the operation of LRFW uncrewed ISR aircraft, the Soviet Union was working to emulate what the United States had accomplished and, more generally, bolster its decidedly inferior airborne ISR capabilities. After all, the Soviet Union had no analogs to the American crewed U-2 and SR-71 and more generally lacked the United States' extensive array of crewed ISR aircraft of various types. The massive supersonic Tu-123, which entered military service in the mid-1960s, was very expensive to build and, as a single-use cruise missile-like uncrewed aircraft, was extremely expensive to operate. Soviet designers therefore worked on the development of successor reusable uncrewed aircraft designs that had much in common with the concept established by the American AQM-34.

Soviet developmental efforts resulted in two new LRFW uncrewed ISR aircraft that entered service by the late 1970s. The first, the Tu-141, weighed over 6,000 kilograms, or around two-thirds of the maximum take-off weight of a fully loaded crewed

MiG-21 fighter, and was heavier and larger than any of the American uncrewed ISR aircraft that preceded it.[53] Powered by a turbojet engine, this subsonic aircraft was designed to undertake autonomous ISR missions at distances of several hundred kilometers from the launch area and as, such, offered the Soviet military broadly comparable ISR capabilities to that offered by some variants of the American AQM-34. The primary distinction between the Soviet and American designs, however, was that the Tu-141 was exclusively launched from the ground (at zero altitude and zero speed) with the assistance of a solid rocket booster and was thereafter recovered through the deployment of a parachute.[54] The Soviet Union also developed a conceptually similar but much smaller design, the Tu-143, which weighed around 1,000 kilograms and had a maximum range of around 200 kilometers.[55]

The Tu-141 and Tu-143 are primarily relevant to the present purposes as examples of the adoption (i.e., emulation), however delayed, of advances in LRFW uncrewed aircraft technology by other states, a dynamic that can moderate the effects of technological change by preventing the monopolization of capability gains to a single state. It is nevertheless important to recognize that the Soviet Union did not develop and deploy the Tu-141 and Tu-143 to substitute for its crewed ISR aircraft and continued to heavily rely on reconnaissance variants of crewed combat aircraft as well as ISR satellites. In other words, the Tu-141 and Tu-143 LRFW uncrewed ISR aircraft, which were never fielded in very large numbers, were not intended to supplant crewed ISR

[53] Jonathan Pike, “Tu-141 Strizh / Tu-143 "Reis,” Global Security, No Date, https://www.globalsecurity.org/military/world/ukraine/tu-141-ssm.htm
[54] The AQM-34 family of aircraft were typically launched from the wings of DC-130 transport aircraft to maximize range and were typically recovered in the air by a helicopter after the deployment of a parachute.
[55] Jonathan Pike, “Tu-143 Reys,” Global Security, No Date, https://www.globalsecurity.org/military/world/russia/tu-143.htm

aircraft but to complement the Soviet Union's crewed ISR aircraft and uncrewed ISR satellites. In this respect, the Soviet experience parallels the American experience, and both experiences provide evidence of the enduring limitations of uncrewed aircraft technology and the competitiveness, or lack thereof, of uncrewed aircraft in so-called tactical reconnaissance roles.

14.3 The Enduring Limitations of LRFW Uncrewed Aircraft Technology

While the likes of the AQM-34, Tu-141, and Tu-143 did not put a human pilot at risk, these uncrewed aircraft had airframes that balanced (subsonic) speed and range instead of an airframe that was optimized for high-altitude operations (as in a crewed U-2), speed (as in the crewed A-12/SR-71 and uncrewed D-21), or endurance (which requires adequate fuel, efficient engines, and high aspect ratio wings offering excellent glide performance). In the United States and the Soviet Union, as in other countries, uncrewed aircraft in the vein of the AQM-34, Tu-141, and Tu-143 were competing for roles and funding with reconnaissance variants of crewed combat aircraft (primarily versions of fighter aircraft from the 1970s onward).

Although the operation of these crewed aircraft, of which the widely employed American (lighter-weight) RF-5 and (heavier-weight) RF-4 are useful representatives, inherently put human operators at risk of harm by air defenses, the underlying aircraft designs were known quantities, could readily integrate into a strike package of crewed combat aircraft (i.e., take photographs of potential targets before an air strike and fly behind a formation of attacking aircraft to conduct bomb damage assessment after an air strike), did not have to emit radio signals that were susceptible to interception and jamming (i.e., the pilot

could operate while undertaking radio silence), could operate in the same meteorological conditions as the strike package, and, not least, were equipped, through the devices of the onboard human operator(s), with the means for independent navigation and adaptive maneuvering in response to changing circumstances so as to evade air defenses, terrain, and inclement weather during both ingress and egress.

To compete for roles and funding, uncrewed aircraft technology had to credibly compete with crewed aircraft and the likes of the Tu-141, Tu-143, and AQM-147 were, notwithstanding their novelty, not very competitive when compared with the RF-5, RF-4, and international analogs of these crewed "tactical" ISR aircraft. As a result, these LRFW uncrewed aircraft were primarily complements to, not substitutes for, extant crewed ISR aircraft and this limited the practical implications of advances in LRFW uncrewed aircraft technology on fielded military capabilities. For this situation to change, uncrewed aircraft technology would have to become more competitive, and this required substantial breakthroughs in the areas of aircraft automation, sensors, and radio technology.

14.4 Background Technological Developments

Before proceeding to examine the implications of advances in uncrewed aircraft technology in the ca. 1970-1990 timeframe, it is productive to pause and briefly examine what are perhaps best characterized as background technological advances made by the 1970s and during the 1970s and 1980s. By background technological advances, I refer to technological developments that affected the realm of possibilities for uncrewed aircraft technology even though these technological advances were typically pursued for unrelated civilian and/or military purposes.

Setting aside nuclear weapons, perhaps the most consequential area of technological development of the post-war period began with the invention of the transistor, the first practical form of solid-state electronics.[56] Before this technological breakthrough, which was made in 1947, electronic systems utilized vacuum tubes, also known as thermionic valves, and the nature of vacuum tube technology, which had been standard since the early twentieth century, resulted in large, heavy, delicate, and inefficient electronic devices. With the advent of solid-state transistors, smaller and lighter electronic devices became a possibility, but this did not transpire overnight and the practical applications, and as such the immediate implications for fielded military capabilities, of this otherwise seminal technological advance were limited. With the development and testing of new forms of transistors, particularly MOSFET transistors, throughout the 1960s, however, successive advances in solid-state electronics facilitated the continuous miniaturization of electronics.[57] Importantly, miniaturization helped to reduce the cost of electronic devices, a dynamic perhaps best understood in terms of the rapid adoption of (relatively) inexpensive personal electronics ranging from walkie-talkies to personal computers. Much as Interwar advances in radio technology facilitated the growing ubiquity of radios in civilian and military roles by the late 1930s, advances in solid-state electronics enabled the growing ubiquity of a wide array of electronic devices over the 1970s and 1980s.

Although a comprehensive account of the implications of advances in solid-state electronics for military technology is well beyond the scope of both this chapter and this

[56] Michael Riordan and Lillian Hoddeson, *Cystal Fire: The Birth of the Information Age* (New York: W.W. Norton & Company, 1997)
[57] Ross Knox Bassett, *To the Digital Age: Research Labs, Start-up Companies, and the Rise of MOS Technology* (Baltimore, MD: Johns Hopkins University Press, 2007)

dissertation, it is nevertheless important to recognize four specific developments related to advances in solid-state electronics that affected the practicality of LRFW uncrewed aircraft technology in the 1970-1990 timeframe. First, by the early 1970s, militaries were unshackled from the limitations imposed by wet film camera technology (which originated in the mid-nineteenth century) as a result of the development of charge-coupled devices (CCD), a type of integrated circuit that enabled the development of electro-optical sensors – digital cameras. Second, advances in solid-state electronics facilitated the development of full duplex radio links that allowed for live two-way broadcasts at progressively greater bandwidths. The combination of electro-optical sensors and improved radio datalinks allowed for the transmission of still imagery and video, as well as other forms of intelligence data, in real-time, thereby offering the feedback required for the practical remote piloting of uncrewed aircraft. Third, advances in solid-state electronics facilitated the development of increasingly capable computers that were smaller and less expensive to build and operate than earlier digital computers. This growth in compact and increasingly more affordable computation power meant that aircraft, whether civilian or military and whether crewed or uncrewed, could be increasingly automated (with automation understood as existing on a spectrum) and miniaturization meant that small computers could provide equal if not greater capabilities than the larger, heavier, more fragile, and more expensive computers of years past. Fourth, solid-state electronics allowed for the development of practical lasers which, while limited in terms of immediate military applications, would in time allow for the development of lasers for the roles of target designation and range finding (both of which were, not incidentally, relatively low-power roles).

Background technological developments did not, of course, only transpire in electronics, and technological advances elsewhere also affected the realm of possibility for LRFW uncrewed aircraft technology. For example, the Space Race between the United States and the Soviet Union drove unprecedented levels of peacetime investment toward research and development which, beyond achievements in crewed spaceflight, resulted in the development and employment of large numbers of increasingly capable satellites. As explained earlier, uncrewed ISR satellites both complemented crewed ISR aircraft and offered states new ISR capabilities that were otherwise unattainable even with exquisite high-altitude aircraft like the U-2 or SR-71. Whereas the first satellites placed into orbit, above all the Soviet Sputnik, were little more than proofs of concept, by the late 1960s, the civilian and military applications of satellite technology extended far beyond ISR to long-range communications and (satellite) television broadcasting. Advances in satellite technology also led to the development of the American Transit/NAVSAT, the first satellite-based radio navigation system. By 1964, Transit satellites began to support the operations of American nuclear-armed ballistic missile submarines, which required an accurate estimate of their location at sea to effectively employ guided ballistic missiles with any meaningful level of accuracy. In time, civilian and military surface ships would similarly end decades of reliance on ground-based radio navigation systems (for a variety of technical reasons, satellite-based navigation had limited implications for aircraft, whether crewed or uncrewed, until the advent of the American Navstar Global Positioning System in the 1990s).

14.5 The United States Renewed Pursuit of LRFW Uncrewed Aircraft in the 1970s

Having briefly summarized the technological context, we can now return to developmental efforts made by the United States, the country at the forefront of essentially every manifestation of uncrewed aircraft technology as of the early 1970s. While the end of the Vietnam War was followed by the retirement of the AQM-34, AQM-91, and the D-21, it did not result in the wholesale abandonment of LRFW uncrewed aircraft technology by the United States. Rather, the United States undertook two development programs with the aim of developing LRFW uncrewed aircraft that were better suited for use in Central Europe against the Warsaw Pact.[58]

The first of these two most notable American developmental efforts undertaken in the 1970s, the Compass Dwell program, aimed to develop an uncrewed ISR aircraft capable of long-duration operations at 40,000 feet as well as taking off from and landing on a runway. Whatever the qualities of the American AQM-34, its inability to operate from runways, a legacy of its origins as a derivative of uncrewed target aircraft technology, restricted it to an expensive and complex aerial launch operation (conducted from a DC-130 aircraft), and a similarly expensive and complex aerial recovery operation (conducted by a heavy lift helicopter). Putting aside the cost of these complex launch and recovery procedures, this operating model resulted in very low sortie rates (if the aircraft was successfully recovered and not damaged) and required the launch aircraft and the recovery helicopter to operate close to the frontlines if the AQM-34 was to fly to anywhere near its maximum range. If an uncrewed aircraft could operate from runways, the thinking went, operating costs could decrease while sortie rates could increase and an

[58] Ehrhard, *Air Force UAVs.*

uncrewed aircraft could therefore better compete with crewed ISR aircraft, such as the American RF-4 (the reconnaissance version of the F-4 Phantom fighter), for funding and roles.[59]

The product of the Compass Dwell program, the GQM-93,[60] used a modified glider (i.e., sailplane) design and was notably an optionally crewed aircraft – it could be flown by a pilot if needed (typically for ferry flights) but was designed for operational use without a pilot onboard. The use of a powered version of a glider airframe was significant but not without precedent. During the Vietnam War, the U.S. military operated a small number of modified gliders (which were converted into powered motor gliders) for use as a crewed ISR aircraft undertaking low-altitude nighttime operations with infrared sensors.[61] By their very nature, gliders, including motor gliders, are fixed-wing aircraft optimized for a high glide ratio and achieve this through unswept high aspect ratio wings. Crewed combat fighter aircraft, in contrast, have airframes optimized for speed and maneuverability and therefore have swept low aspect ratio wings. The AQM-34, having been developed from an uncrewed target aircraft designed to serve as a surrogate for crewed fighter and bomber aircraft and carried on the wings of a DC-130 cargo aircraft, had a design that was similarly optimized for speed and maneuverability instead of glide performance.

Although it was not intended to supplant the crewed U-2 and in any event lacked the maximum ceiling, speed, and payload capacity to serve as an equally capable

[59] Ehrhard, *Air Force UAVs.*
[60] "E-Systems GQM-93," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-93.html
[61] "Lockheed YO-3A Quiet Star," The Museum of Flight, No Date, https://www.museumofflight.org/exhibits-and-events/aircraft/lockheed-yo-3a-quiet-star ; Norman Polmar, "Historic Aircraft – Silent (By) Night," *Naval History Magazine* 31, No. 5 (2017), https://www.usni.org/magazines/naval-history-magazine/2017/october/historic-aircraft-silent-night

substitute for the U-2, the GQM-93 had a major limitation that affected its suitability for peacetime operations over Central Europe. Given the density of peacetime civilian air traffic above West Germany and adjacent countries in peacetime, the operational context for which the GQM-93, an aircraft inherently vulnerable to air defenses, was developed, the designers of the GQM-93 were required to develop an aircraft that could operate at an altitude of over 40,000 feet. The inability of the aircraft to meet this requirement resulted in the cancellation of the Compass Dwell program in 1973, but this was followed by another, more ambitious American developmental effort, the Compass Cope program.

Given the United States' extensive array of airborne ISR capabilities, against which all proposals for an LFW uncrewed ISR aircraft would compete, the design objectives for the Compass Cope program focused on developing a high-altitude aircraft capable of undertaking persistent (i.e., long endurance) electronic surveillance flights over Central Europe in order to detect Warsaw Pact air defenses through their radar and radio emissions. The United States means of detecting Warsaw Pact air defenses in the form of electronic intelligence satellites and the crewed RC-135, an electronic reconnaissance version of what is essentially a Boeing 707 airliner, among other aircraft. Notwithstanding its many qualities, the RC-135's airframe was incapable of high-altitude operations, which extends the line-of-sight of onboard sensors, and human physiological limitations precluded the large crew inside the cramped aircraft (packed with electronics) from undertaking long endurance missions. As a result, the RC-135 was far from an ideal platform to undertake persistent surveillance missions at a time when Warsaw Pact forces were increasingly operating more mobile surface-to-air missile systems that were not tied to fixed prepared sites in the manner of earlier Soviet air defense systems. In other words,

the Compass Cope program attempted to develop an LRFW uncrewed aircraft that could do what existing crewed aircraft could not.

By 1974, the two prototype designs competing in the Compass Cope program, the YQM-94 and YQM-98, had both undertaken test flights. These uncrewed aircraft operated from runways like crewed fixed-wing aircraft, were equipped with more efficient turbofan engines, weighed over 6,000 kilograms, and, not unimportantly, had high aspect ratio wings and large wingspans of around 25 meters.[62] By way of comparison, the highly maneuverable F-5E, a supersonic light fighter with a maximum takeoff weight of around 11,000 kilograms, had a wingspan of just over 8 meters, while a 727-100, a 120-seat narrow-body airliner with a maximum takeoff weight of around 77,000 kilograms, had a wingspan of around 33 meters.

By any measure, the YQM-94 and YQM-98 were large and heavy aircraft with very large wingspans and constituted ambitious efforts to develop LRFW uncrewed aircraft that could compete with crewed ISR aircraft in terms of performance. It is nevertheless important to recognize that these were both remotely piloted aircraft controlled in real time over a radio link with very limited onboard automation that practically relegated these aircraft to peacetime operations in the central European military-geographical context for which these were intended. Ultimately, a combination of cost overruns and the loss of expensive prototypes in crashes brought developmental efforts to a close without either of the competing prototypes being ready for service.[63] As

[62] "Boeing GQM-94 B-Gull," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-94.html; "Teledyne Ryan GQM-98 R-Tern," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-98.html
[63] Ehrhard, *Air Force UAVs.*

a result, these expensive developmental efforts engendered negligible to non-existent effects on fielded military capabilities.

Had the YQM-94 and YQM-98 been designed by a hypothetical state that, for whatever reason, possessed the resources to undertake such costly and complex developmental efforts but either did not operate crewed ISR aircraft capable of undertaking the same roles or was incapable of developing (or importing) such a crewed ISR aircraft, then it is possible that said hypothetical state would have doubled down and pursued the Compass Cope program or a successor developmental effort until it bore fruit in the form of a functional, service-ready design. In the real world, however, the YQM-94 and YQM-98, which were by a wide margin the most advanced LRFW uncrewed aircraft designs then in existence, were developed by a state that had the necessary resources to pursue a crewed aircraft design in the event that the development of the uncrewed aircraft did not work out. Tellingly, the cancellation of the Compass Cope program was followed by the United States' development and manufacture of a new version of the crewed U-2 aircraft equipped with the sensors required to undertake high-altitude electronic intelligence missions.

While the latest version of the crewed U-2, the TR-1 (later redesignated as the U-2S), which had a maximum endurance of 12 hours, could not match the endurance offered by the uncrewed YQM-94 and YQM-98, it was a proven and reliable multi-role (in terms of various ISR roles) aircraft capable of operating at higher altitudes than either of the uncrewed prototype designs. It also had greater payload capacity, an important consideration given the weight and volume of high-end sensors required for high-altitude

operations that could take advantage of the extended radar and radio horizon.[64] Most importantly, the onboard human pilot of the U-2, augmented by advances in autopilot and navigation technology, could independently undertake complex missions and adapt to changing environments in a manner that was far beyond the reach of even the latest advances in aircraft automation technology.

The experience of the American Compass Dwell and Compass Cope programs highlights the enduring limitations of LRFW uncrewed aircraft technology as of the mid to late 1970s as well as the recurring theme that uncrewed aircraft had to be competitive against crewed aircraft designs to obtain funding and roles in a world of finite resources, pressing operational requirements, and continuous advances in crewed aircraft technology. While the advent of solid-state electronics facilitated major advances in uncrewed aircraft technology, above all in the areas of sensors, datalinks, and onboard automation, these background technological advances also enabled growth in the capabilities offered by the latest crewed military aircraft. If LRFW uncrewed aircraft technology was to offer states significant military advantages that justified the cost of developing a new military technology at great expense and putting a new design into production, it would have to offer something that existing and prospective crewed aircraft technology could not.

[64] "Senior Year / Aquatone / U-2 / TR-1," Federation of American Scientists, No Date, https://irp.fas.org/program/collect/u-2.htm

14.6 Israel's Development of LRFW Uncrewed Aircraft In the 1970s and 1980s

The United States, the undisputed leader in uncrewed aircraft technology as of the mid-1970s, enjoyed extensive resources and fielded a large and capable military that was equipped with a large and diverse array of crewed aircraft that fulfilled virtually every conceivable role. As a result, from the 1960s onward, the United States primarily pursued large fixed-wing uncrewed aircraft for missions that its crewed aircraft – and the crewed aircraft of all other countries – either could not undertake or were unsuited to undertake. For the United States, this meant developing large fixed-wing uncrewed aircraft to undertake high-risk missions over defended airspace or long-endurance ISR roles that exceeded human physiological limitations. The United States was not, however, the only country making substantive advances in LRFW uncrewed aircraft technology in the ca. 1970-1990 timeframe, and other countries experienced different circumstances and therefore pursued uncrewed aircraft technology to satisfy their own often disparate requirements.

The circumstances that Israel, the other pioneer in LRFW uncrewed aircraft in this timeframe, found itself in the 1970s were very different than that of the United States, and the interplay of Israel's unique – among major military powers – military-geographical context and military requirements, and new possibilities enabled by background technological developments, facilitated its development of a *new type* (or category) of LRFW uncrewed aircraft. While the immediate implications in terms of fielded military capabilities were limited, Israel's efforts established a new direction for LRFW uncrewed aircraft technology that would have greater implications for fielded military capabilities in the long run.

Like all advanced militaries in the post-1945 period, Israel operated uncrewed aircraft to serve as targets in support of the training of fighter pilots and ground-based air defenses. During the 1973 Arab-Israeli War, Israel initially experienced unsustainably high losses of crewed combat aircraft at the hands of the latest Soviet air defense systems operated by Egypt and Syria.[65] In response, Israel employed some of its turbojet-powered uncrewed target aircraft, which were designed to serve as realistic surrogates for crewed combat aircraft, in the role of decoys that would force the activation of air defense systems that could in turn be targeted by Israel's crewed combat aircraft.[66] This ad hoc employment of uncrewed target aircraft in this role was innovative but far from unprecedented; the United States had, after all, employed some of its AQM-34 uncrewed aircraft in decoy roles during the Vietnam War.[67] Nevertheless, this experience, one in which uncrewed aircraft were employed in a single-use role to augment crewed combat aircraft, highlighted the potential applications of large and otherwise reusable fixed-wing uncrewed aircraft technology.

Beyond the challenges posed by the latest in Soviet air defense technology, the 1973 war exposed a variety of shortcomings in Israeli military capabilities including inadequate ISR capabilities over the battlefield. Although Israel's military typically operated in sparsely vegetated environments and in atmospheric conditions characterized by little in the way of cloud cover that was highly conducive to effective aerial reconnaissance operations, Israeli ground forces repeatedly struggled to identify enemy

---

[65] Abraham R. Wagner, *The Lessons of Modern War: Volume 1: The Arab-Israeli Conflicts, 1973–1989* (Boulder, CO: Westview, 1990).; Abraham R. Wagner, *The Lessons of Modern War: Volume 1: The Arab-Israeli Conflicts, 1973–1989* (Boulder, COL: Westview, 1990).

[66] John F. Kreis, "Unmanned Aircraft in Israeli Air Operations," *Air Power History* 37, No. 4 (1990): 46-50.

[67] Kreis, "Unmanned Aircraft in Israeli Air Operations."

forces before encountering enemy fire. The Israeli military operated reconnaissance variants of crewed combat aircraft but these were few in number and in high demand undertaking a variety of priority missions in support of both air and land operations. Whereas the United States and leading militaries worldwide increasingly bolstered their battlefield ISR capabilities through the deployment of reconnaissance and attack helicopters, the Israeli military of 1973 only operated transport helicopters that were not suitable for this role, particularly in a sparsely vegetated environment that offered few opportunities for concealment at low altitudes and behind tree cover.

Lacking the resources required to procure and operate fixed- and rotary-wing crewed ISR aircraft in large numbers, Israel developed a series of prototype large reusable fixed-wing uncrewed aircraft designs that took advantage of the recent background advances in computing (aircraft automation), sensors (electro-optical sensors), and communications (full-duplex datalinks) technology. The two most notable early Israeli designs, which entered service in the late 1970s as the IAI Scout and the Tadiran Mastiff, were broadly similar in terms of specifications. These uncrewed aircraft weighed no more than 150 kilograms, had wingspans no greater than 5 meters, and could be equipped with a maximum payload of just 40 kilograms. Equipped with small piston engines that facilitated a maximum speed no greater than 180 km/h, these propeller-driven uncrewed aircraft were closer in terms of speed and payload to (crewed) light helicopters and small (civilian) piston-engine crewed aircraft, both of which, it is important to emphasize, could be used to carry the very same sensors and datalinks.

Compared with the latest in then-contemporary fixed- and rotary-wing military aircraft, the Scout and Mastiff were unimpressive and, in many respects, had much in

common with large remote-controlled hobby aircraft, albeit hobby aircraft that benefitted from advances in solid-state electronics. What the Scout and Mastiff lacked in terms of payload and speed, however, was outweighed by a maximum endurance of over six hours, something that most crewed aircraft, particularly smaller and less expensive designs, could not match. Moreover, the Scout and Mastiff offered this endurance with airframes – excluding the onboard computers, sensors, and communications systems – that cost substantially less to procure and operate than crewed ISR aircraft and, no less importantly, could undertake battlefield ISR missions without putting any human operators at risk. Stated differently, advances in solid-state electronics provided Israeli designers with suitable autopilots, electro-optical sensors, and datalinks that allowed for the development of miniaturized uncrewed long endurance aircraft that were not only suitable for the battlefield ISR role but could also partially substitute for more expensive fixed- and rotary-wing aircraft in this role.

14.7 Viewing Israeli Technological Developments in their Proper Context

On paper, the likes of the Mastiff and Scout would seem to suggest that Israel uncovered a winning formula for obtaining battlefield ISR capabilities, one that all militaries should have rushed to adopt. As explained throughout this dissertation, however, the effectiveness, and as such the implications, of all military technologies are coproduced with the actions taken and forgone by adversaries and are shaped by the military-geographical context for which a new military technology is developed and in which it is operated.

To understand why Israel, and not another country, was the first to develop and field miniaturized LRFW uncrewed aircraft in the battlefield ISR role, we need to give greater consideration to Israel's military-geographical context and compare it with the military-geographical contexts that were the central focus of other leading militaries in this timeframe. As the multiple American developmental efforts covered thus far in this and the prior chapter highlight, less permissive military-geographical contexts can increase military requirements such that an uncrewed aircraft cannot productively compete with a crewed aircraft for funding and roles. The logical corollary is that more permissive military-geographical contexts can reduce military requirements such that an uncrewed aircraft can more productively compete with a crewed aircraft for funding and roles. If the Soviet Union and its allies did not operate air defenses, for example, then not only would the crewed U-2 have remained viable for overflights, but uncrewed aircraft such as the AQM-34 could have plausibly undertaken the same role. Stated differently, it is entirely plausible that variant military-geographical contexts may lead one state to pursue and productively employ a newly available technology that another considers to be impractical without either state necessarily making an error of judgment. In the following paragraphs, I will offer three non-exhaustive ways in which Israel's military-geographical context was more considerably more permissive to the operation of early miniaturized LRFW uncrewed ISR aircraft such as the Scout and the Mastiff than those encountered by other leading militaries including the United States and the Soviet Union.

First, the Israeli Scout and Mastiff primarily operated in areas that were much unlike Southeast Asia or Central Europe and instead characterized by limited cloud cover and infrequent inclement weather, thereby facilitating regular high visibility flying

conditions. In both Southeast Asia and Central Europe, fog, low cloud cover, and precipitation regularly interfered with the operations of crewed aircraft. Remotely piloted LRFW uncrewed aircraft like the Israeli Mastiff and Scout were no better positioned to operate in such flying conditions as remote pilots had inferior sensory acuity, more limited control feedback than an onboard pilot, and more limited reference to visual cues. Israeli designers could, however, make do with LRFW aircraft that were restricted to operations in high visibility flying conditions (in daylight until suitable infrared sensors became available) as this was the military-geographical context in which Israeli aircraft, crewed and uncrewed, regularly operated.

Second, the interplay of typically high visibility at flight altitude and the sparsely vegetated environment found along most of Israel's frontiers meant that Israel's miniaturized uncrewed aircraft were better positioned to detect objects with what were by any standard small, light, and low-resolution electro-optical sensors. Higher resolution electro-optical sensors were available at the time but were far larger and heavier than either the Mastiff or Scout could carry (and were also far more expensive). It is important to recognize that electro-optical sensors, like human eyes and other electromagnetic sensors, are inherently limited in terms of field of view and the capacity to resolve an object from its surroundings. Equipped with a telephoto lens, the electro-optical sensor on an uncrewed aircraft could zoom into a specific area (at the cost of not observing any other areas), but all things equal, the more complex the background in which an object of interest is situated, the greater the challenge of resolving it from surrounding objects in its background. As a result, the remote operator of an uncrewed aircraft observing a low-resolution live video feed is better positioned to detect and track objects of interest in a

high-visibility environment with sparse vegetation than in a more variable visibility environment with denser vegetation. Stated differently, a Mastiff or Scout would be more practical and effective as ISR aircraft along Israel's frontiers than in Southeast Asia or Central Europe.

Third, although Israel's adversaries operated air defense systems, the air defense environment along Israel's frontiers was very different from that found in Central Europe and Israel's adversaries did not deploy extensive electronic warfare capabilities that could be used to jam the communications link through which remotely piloted aircraft both operated and transmitted sensor data. To appreciate the extent to which Israel was in an increasingly preponderant position in terms of air combat capabilities by the late 1970s and early 1980s, one need not look further than the manner in which Israeli aircraft, crewed and uncrewed, operated with near impunity outside of Israeli airspace during both "peacetime" and in combat operations. In Central Europe, neither NATO nor the Warsaw Pact, out of shared concern for escalation, whether conventional or nuclear, dared to operate aircraft, crewed or uncrewed, in such an aggressive manner inside the other side's airspace during peacetime (even during crises) and the definition of a practical uncrewed aircraft capable of substituting for crewed aircraft in that specific military-geographical context entailed an aircraft capable of credibly operating in wartime, in an environment in which air defenses were fully activated and extensive electronic warfare operations, both friendly and hostile, were underway. This, of course, constituted a far more challenging set of technical requirements to meet than those faced by Israel.

By the early 1980s, Israel, already operating in a military-geographical context that was highly permissive to the operation of miniaturized LRFW aircraft in the vein of

the Mastiff and Scout, was in a much stronger military position than it had been during the 1973 war. The 1978 peace accords with Egypt stabilized Israel's western frontier (Israeli forces did not fully vacate the Sinai Peninsula until 1982) and Israel's air combat capabilities increased considerably as a result of extensive procurement and military aid from the United States. By the early 1980s, Israel had one of the most capable air forces in the world, one equipped with scores of some of the latest American crewed combat aircraft designs such as the F-15 and F-16 that Israel's adversaries could not hope to match. Such was the extent of Israel's qualitative advantage in technology and non-material factors, at least in terms of air combat capabilities, that Israel successfully undertook a complex and high-risk long-range operation to attack a nuclear reactor under construction in Iraq in 1981 and, in 1982, imposed a crushing defeat on the Syrian air force. The extent of Israel's military advantages over its adversaries is important for the present purposes because (A) relatively small numbers of miniaturized LRFW uncrewed ISR aircraft were in Israeli service as of the early 1980s and (B) the primary evidence for the military implications of LRFW uncrewed aircraft in the ca. 1970-1990 timeframe – excluding U.S. operations over Vietnam in 1972 – is found in Israel's employment of its early LRFW uncrewed ISR aircraft against Syrian forces in the early 1980s and Israel's experience is often highlighted in writings that characterize advances in LRFW uncrewed aircraft technology as engendering major shifts in military capabilities.

14.8 The Gradual Adoption of Miniaturized LRFW Uncrewed Aircraft

While most leading militaries did not operate in as permissive a military-geographical context as that which the Israeli military increasingly found itself in from the late 1970s

onwards, Israel's successful employment of the new category of miniaturized LRFW uncrewed aircraft in the early 1980s did not go unnoticed and was followed by the moderating process of adoption. By the mid-1980s, leading militaries were responding to the deployment of growing quantities of new models of equipment, whether tanks, aircraft, or submarines, the result of both converging modernization cycles and increased diplomatic tensions, and were increasingly concerned with the implications of ongoing advances in the (other) area of uncrewed aircraft technology better known as precision-guided munitions.

From the mid-1970s onwards, the United States and NATO sought to harness the capability gains offered by the military applications of advances in solid-state electronics to counter the quantitative superiority of Warsaw Pact ground forces. These efforts resulted in the American operational concept of AirLand Battle and the closely related NATO concept of Follow-on Forces Attack, both of which required major improvements in battlefield ISR capabilities to implement. The Soviet Union, for its part, was well aware of these developments and influential segments of the Soviet military intellectual community were apprehensive of the potential implications of the pairing of improved battlefield ISR capabilities and guided munitions long before such developments had come to pass.[68] In this context, leading militaries made large-scale investments toward improving their battlefield ISR capabilities. Crewed aircraft of all types increasingly received the latest in sensor and datalink technology including sensors of far greater

[68] Dima P. Adamsky "Through the Looking Glass: The Soviet Military-Technical Revolution and the American Revolution in Military Affairs," *Journal of Strategic Studies* 31, No. 2 (2008): 257–94.; Vladimir Kontorovich, *The Impoverished Superpower: Perestroika and the Burden of Soviet Military Spending* (San Fransisco, CA: Institute for Contemporary Studies Press, 1990).; Marc Trachtenberg, "Assessing Soviet Economic Performance During the Cold War: A Failure of Intelligence?" *Texas National Security Review* 1, No. 2 (2018): 76-101.

sophistication than those used on miniaturized LRFW uncrewed aircraft like the Israeli Mastiff and Scout, both of which could only carry a small payload weighing no more than 50 kilograms.

In terms of LRFW uncrewed ISR aircraft, Canada, West Germany, and France procured the Canadair CL-289,[69] a turbojet-powered uncrewed ISR aircraft which, while a further development of the earlier CL-89, incorporated the latest in sensor and datalink technology. Whereas the Israeli Mastiff and Scout were designed to loiter in a rather permissive military-geographical context, the CL-289 was designed to substitute for the reconnaissance versions of crewed combat aircraft to provide updated aerial imagery to the commanders of large ground formations in the Central European operating environment. The United Kingdom, for its part, emulated the Israeli approach and developed a propeller propeller-driven aircraft, the BAE Phoenix (British forces notably operated in the North German Plain),[70] while the United States undertook a series of developmental efforts of its own, which will be examined in the following pages, and later fielded an Israeli design as the RQ-2 Pioneer.[71] Even the Soviet Union, which was decidedly lagging across electronics technology, worked to develop a miniaturized LRFW uncrewed ISR aircraft and deployed at least one system, the Yakovlev Pchela, in a limited capacity, while also upgrading its existing ground-launched, parachute-recovered

[69] "CL-289/ AN/USD-502," Federation of American Scientists, No Date, https://man.fas.org/dod-101/sys/ac/row/cl-289.htm
[70] "Phoenix Battlefield Surveillance UAV," Army Technology, June 8, 2000, https://www.army-technology.com/projects/phoenixuav/?cf-view
[71] "Pioneer Short Range Drone," Federation of American Scientists, No Date, https://irp.fas.org/program/collect/pioneer.htm; "Pioneer RQ-2A UAV," National Air and Space Museum, No Date, https://airandspace.si.edu/collection-objects/pioneer-rq-2a-uav/nasm_A20000794000

turbojet-powered Tu-143 with electro-optical sensors and a datalink so that it could share sensor data in real-time.[72]

14.9 The United States Renewed Pursuit of LRFW Uncrewed ISR Aircraft in the 1980s

While Israel's efforts made the United States a laggard in the development and deployment of miniaturized LRFW uncrewed aircraft in the early 1980s, the country returned to the forefront of uncrewed aircraft technology by the end of the decade. Some of these efforts are worth recounting here not only as evidence of the speed at which a pioneering country, Israel, could lose its technological lead, such as it was, but also because the American experience in the 1980s highlights the reality that technological advances had not created a situation in which even a country with the resources and engineering capacity of the United States could open the proverbial tap and receive an outpouring of service-ready LRFW uncrewed aircraft designs that could effectively compete with crewed aircraft for funding and roles. Technological development remained protracted, and this moderated whatever gains in fielded military capabilities that advances in LRFW uncrewed aircraft technology could offer.

Israel's well-publicized employment of miniaturized LRFW uncrewed aircraft to locate Syrian mobile surface-to-air missile systems in 1982,[73] followed as it was by the destruction of said missile systems with primarily American-built munitions launched

[72] Kreis, "Unmanned Aircraft in Israeli Air Operation."
[73] Haim Yogev, Ronen A. Cohen, and Eyal Lewin, "Revolution in Military Affairs – The Operation Mole Cricket as a Case Study for the Technological Race During the Cold War," *International Area Studies Review* 25, No. 2 (2022): 138-156.

from American-built crewed combat aircraft, unsurprisingly did not escape the notice of the American military. By the mid-1980s, the United States was focused on improving its capacity to target the latest, more mobile Soviet air defense systems, and to this end it started to deploy the AGM-88 HARM, a new autonomous air-launched supersonic missile designed to home in on the emissions of air defense radars.[74] The rocket-powered AGM-88 was the replacement for the Vietnam War-era AGM-45 Shrike and AGM-78 Standard and offered many advantages relative to these older systems but it was wholly optimized to reduce time to target in order to damage or destroy a detected emitting mobile surface-to-air missile system before it could relocate. The new missile was, as such, unsuited to detecting, let alone engaging, non-emitting mobile surface-to-air missile systems in the manner that a miniaturized LRFW uncrewed ISR aircraft equipped with an electro-optical sensor, such as those employed by Israel in 1982, could (without, of course, any ability to attack a detected surface-to-air missile system).

To better counter the latest Soviet mobile air defense systems, the United States undertook an ambitious program to develop the subsonic AGM-136 Tacit Rainbow,[75] an air-launched fixed-wing subsonic uncrewed aircraft that blurred the (artificial) distinction commonly drawn between cruise missiles, as a form of guided munition, and large fixed-wing uncrewed aircraft. The AGM-136 is notable for the present purposes because it was designed to follow waypoints to a target area several hundred kilometers from its launch point – in the same manner as contemporary cruise missiles and earlier LRFW uncrewed

[74] "Raytheon (Texas Instruments) AGM-88 HARM," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-88.html; "AGM-88 HARM (High Speed Antiradiation Missile)," Federation of American Scientists, No Date, http://man.fas.org/dod-101/sys/smart/agm-88.htm
[75] "AGM-136 Tacit Rainbow," Federation of American Scientists, No Date, https://man.fas.org/dod-101/sys/smart/agm-136.htm; "Northrop AGM/BGM-136 Tacit Rainbow," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-136.html

ISR aircraft such as the AQM-34 – and thereafter loiter over a designated area until it detected an actively emitting radar within range of the onboard passive radiation homing sensor, after which the AGM-136 would autonomously navigate toward and detonate on impact with the detected radar without any human intervention (i.e., no human in the loop). What the AGM-136 offered was, in other words, a means of using large (single-use) fixed-wing uncrewed aircraft to substitute for crewed combat aircraft in an increasingly dangerous role behind enemy lines. Importantly, the AGM-136 was not intended to supplant the AGM-88 HARM or the crewed combat aircraft that employed this supersonic missile; the capacity to loiter was only relevant for operations behind frontlines as the United States had extensive capabilities, including crewed combat aircraft equipped with the AGM-88, to detect and target Warsaw Pact mobile air defense systems situated close to the front lines.

The United States similarly pursued the development of a ground-launched miniaturized LRFW uncrewed aircraft, the MQM-105 Aquila,[76] to loiter over the battlefield to detect and track moving targets. Just as the AGM-136 was developed to operate behind frontlines, the MQM-105, an aircraft in the same weight class as the Israeli Mastiff and Scout, was developed to support "deep fires" by facilitating the operation of the M712 Copperhead, a 155 mm artillery shell equipped with a semi-active laser homing (SALH) guidance system.[77] On paper, the guided M712 artillery shell, which is itself a manifestation of (unpowered) uncrewed aircraft technology given that SALH guidance is essentially a form of remote control, offered American artillery units

[76] "Lockheed MQM-105 Aquila," Designation Systems, No Date, https://www.designation-systems.net/dusrm/m-105.html
[77] "M712 Copperhead," Federation of American Scientists, No Date, https://man.fas.org/dod-101/sys/land/m712.htm

the ability to undertake precise attacks on moving targets, including tanks, over distances of around ten kilometers and was, as such, a core element of the aspirations encapsulated in the American operating concept of AirLand Battle. While artillery units could rely upon their ground-based forward observers to designate – "paint" – a target with a laser designator that the M172's onboard guidance system would acquire in flight, (large numbers of) human forward observers could not realistically operate far beyond the frontline in the Central European military-geographical context given the high force densities along the frontlines. In principle, crewed fixed- and rotary-wing aircraft could designate targets for artillery units, but this was increasingly unrealistic in the air defense and air combat environment that characterized the Inner German border by the 1980s, and artillery units required freshly updated target location data to engage Warsaw Pact ground forces moving and attacking in echelons. Through the employment of an LRFW uncrewed aircraft like Aquila, which was equipped with electro-optical sensors, a datalink for real-time transmission of sensor data, and a laser designator, artillery units equipped with the laser-guided M172 artillery round could, in principle, engage targets behind the front lines without putting any human pilots in harm's way and thereby support the large-scale employment of ground-launched guided munitions.

The third and most ambitious American developmental effort resulted in the Boeing Condor, a large prototype aircraft with a maximum take-off weight of around 9,500 kilograms that was the most advanced LRFW uncrewed aircraft at the time of its first flight in 1988.[78] Making extensive of composite materials to reduce the weight of the airframe, the Condor had a 62-meter wingspan, which is to say a wingspan twice as great

[78] Ehrhard, *Air Force UAVs.*

as the crewed U-2 and one slightly longer than that found on the incomparably larger and heavier Boeing 747 long-range passenger aircraft. Powered by two piston engines with a combined output of 350 horsepower, the propeller-driven Condor could operate at high altitudes for over 48 hours, which was both unprecedented and far beyond human physiological endurance. To undertake such high-endurance missions, the Condor was equipped with the latest in aircraft automation technology such that it could independently operate without a human in the loop. Given these performance parameters and its slow operating speed, the Condor was optimized for persistent high-altitude ISR missions over a selected area and was impractical in most other ISR roles, particularly ISR roles that required operating in areas in which the Condor could be targeted by surface-to-air missile systems or crewed combat aircraft. With the end of the Cold War, the Condor program, which was a developmental effort, was canceled and laid the foundations for the development of what became the RQ-4 Global Hawk in the 1990s.[79]

Different though the AGM-136 Tacit Rainbow, MQM-105 Aquila, and the Boeing Condor were as specimens of uncrewed aircraft technology, these developmental efforts are relevant for the present purposes for two reasons. First, these development programs highlight the tendency of the United States, the leader across all areas of uncrewed aircraft technology throughout the Cold War, to pursue uncrewed aircraft technology for roles that crewed aircraft either could not undertake or were not suitable for undertaking (which encompasses a desire to minimize the loss of human operators at least in peacetime). This tendency, one that this chapter has identified as transpiring as early as the 1950s, may reflect some level of organizational (i.e., service-level) or

[79] Ehrhard, *Air Force UAVs.*

national-level preferences toward keeping a human in control, which could suggest that the United States was insufficiently innovative and squandering of its lead in uncrewed aircraft technology, but it is important to recognize that many datapoints are at variance with such an interpretation. Even as the United States pursued uncrewed aircraft technology, broadly defined, to primarily augment rather than supplant crewed aircraft and later pursued LRFW uncrewed aircraft to substitute for crewed aircraft in high-risk roles, the country eagerly and enthusiastically adopted – and retained – ground- and submarine-launched ballistic missiles for the nuclear delivery role and, among other things, eagerly and enthusiastically adopted and retained increasingly automated crewed combat aircraft in which human pilots were primarily in an aircraft to operate combat systems and make adaptive decisions that onboard automation systems could not.

During the 1980s, the period in which the United States was developing a variety of LRFW uncrewed aircraft of all shapes and sizes, the country was also at the forefront of other areas of uncrewed aircraft technology and deployed a wide range of uncrewed aircraft to both complement and substitute for crewed combat aircraft. For example, in 1983, the United States deployed the RGM/UGM-109 Tomahawk cruise missile, a single-use large fixed-wing uncrewed aircraft that, given its high level of (bounded) onboard automation, could have been employed as a reusable ISR aircraft had it been modified to include sensors, a datalink, and a parachute recovery system instead of a large warhead.[80] While the RGM/UGM-109 Tomahawk allowed American surface

[80] While this may appear to be a farfetched idea, in the late 1980s, an American company, Teledyne Ryan – the company that developed the Ryan Firebee, AQM-34, AQM-91, and the YQM-98 – developed the Scarab, a turbojet-powered LRFW uncrewed aircraft similar in function and configuration to a cruise missile that was equipped with camera to take aerial imagery while following a preprogrammed flight path. The Scarab was only procured and operated by one country, Egypt, and was not very competitive for funding and roles in the context of the United States, which had extensive crewed and uncrewed ISR assets,

warships and submarines to partially substitute for crewed combat aircraft in strikes against ground targets (nuclear-armed versions of the Tomahawk were also developed and deployed), the United States was also undertaking the development of a highly automated active radar homing (ARH) air-to-air missile that, unlike earlier semi-active radar homing (SARH) air-to-air missiles, functioned as a fire-and-forget system that did not rely upon the launch aircraft's human-operated radar to continuously illuminate (i.e., "paint") the target aircraft throughout the engagement. This developmental effort, which resulted in the AIM-120 AMRAAM's entry into service in 1991, augmented the capabilities of crewed combat aircraft but marked yet another milestone in the progressive automation of crewed combat aircraft, a dynamic that dramatically reoriented the role of the onboard human pilot to that of a systems operator capable of undertaking the tasks that aircraft automation technology could not.[81] Both the Tomahawk cruise missile and the AIM-120 air-to-air missile allowed the United States military to put specimens of uncrewed aircraft technology in harm's way and thereby reduce the threat posed to the human operators of equipment that would otherwise be required to undertake such roles.

Second, as of the late 1980s, LRFW uncrewed aircraft technology was not at a state of advancement that allowed militaries to open a proverbial tap – allocate resources

---

but nevertheless demonstrates that a highly automated "cruise missile-like" LRFW uncrewed ISR aircraft became very much possible around the same time that a new generation of cruise missiles like the American Tomahawk entered service. Teledyne Ryan would later develop the BQM-145 Peregrine in the 1990s, an air-launched "cruise missile-like" LRFW uncrewed ISR aircraft that blurred the line between a cruise missile and uncrewed aircraft in the same manner as the 1960s era AQM-34 and the 1980s era Ryan Scarab.

[81] See for example Steven A. Fino. *Tiger Check: Automating the US Air Force Fighter Pilot in Air-to-Air Combat, 1950‑1980.* (Baltimore, MD: Johns Hopkins University Press, 2017). For an overview of air force weapons development in this timeframe, see Kenneth P. Werrell, *Chasing the Silver Bullet: U.S. Air Force Weapons Development from Vietnam to Desert Storm* (Washington, D.C.: Smithsonian Institution Scholarly Press, 2003)

– and quickly obtain service-ready designs that were practical, effective, and, not least, competitive with crewed aircraft in a given role. While successive technological advances expanded the realm of technological possibilities and allowed designers to conceive of LRFW uncrewed aircraft capable of operating in a manner that was previously impossible to implement, developmental efforts continued to be protracted and uncertain. Tellingly, none of the three aforementioned developmental efforts undertaken by the United States in the 1980s – the AGM-136 Tacit Rainbow, MQM-105 Aquila, and the Boeing Condor – resulted in an operational system that engendered any substantive shift in the fielded military capabilities of the United States. Instead, these three developmental efforts encountered some combination of technical difficulties, cost overruns, and political opposition as a result of testing failures and/or cost overruns, such that all were canceled either before or immediately following the end of the Cold War,[82] a major development that itself reduced the impetus for the undertaking of expensive and high-risk research and development efforts.

Uncrewed aircraft technology was an area of considerable research and development activity in the ca. 1970-1990 timeframe. Whether the rather crude if innovative small uncrewed ISR aircraft employed by Israel or the much more sophisticated designs developed by the United States, indicators abound of substantial expansions of the technological frontier relative to earlier timeframes. Nevertheless, the shifts in fielded military capabilities engendered by technological change were neither particularly sizeable nor abrupt. In part, this was because there was a mismatch between the requirements of most militaries and what the latest in LRFW uncrewed aircraft could

[82] Ehrhard, *Air Force UAVs.*

offer but qualitative and quantitative considerations were also important. Simply put, the capability gains accrued to Israel in this timeframe by its development and deployment of LRFW uncrewed ISR aircraft were a function of a more permissive military-geographical context that other states did not benefit from. Change the military-geographical context and, as such, change the military requirements, and advances in LRFW uncrewed aircraft technology in the ca. 1970-1990 had far more subdued military effects. More generally, whatever the qualities of LRFW uncrewed aircraft and the nominal monadic capability gains these offered states, very few specimens existed as late as 1990, and the military effects therefore lagged in purely quantitative terms. There is, as such, very qualified empirical support to be found in the area of uncrewed aircraft technology in the ca. 1970-1990 timeframe in support of the proposition that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states – it very much depends on the state, its military-geographical context, the capabilities of other states, and how many specimens of a new technology are produced.

Chapter 15: Uncrewed Aircraft Technology ca. 1990-2010

The ca. 1970-1990 timeframe provides many examples in which the technological frontier was considerably expanded without major shifts being engendered in the fielded military capabilities of states. While Israel benefitted – enjoyed an increase in its fielded military capabilities – by virtue of being able to employ relatively basic LRFW uncrewed aircraft in a more permissive military-geographical context, other states did not. It is however possible that the previous chapter's temporal coverage was too restrictive and that extending the longitudinal analysis to the ca. 1990-2010 timeframe, in which LRFW uncrewed aircraft technology was further developed, more widely fielded, and progressively diffused, will provide evidence in support of the proposition that technological change tends to engender sizeable and abrupt shifts in the fielded military capabilities of states.

As the contrasting experiences of Israel and the United States in the ca. 1970-1990 timeframe highlight, new military technologies are neither developed nor deployed in a vacuum. It is, therefore, important to consider both supply-side and demand-side dynamics as well as the military-geographical context for which a new military technology is intended when considering the implications of technological change on the fielded military capabilities of states. While sustained technological advances expanded the realm of possibility for LRFW uncrewed aircraft technology in the ca. 1990-2010 timeframe, the end of the Cold War and the primarily lower intensity conflicts that followed, most of which were characterized by extreme asymmetries in air combat and air defense capabilities, affected both demand-side and supply-side dynamics. Simply put, the new post-Cold War military-geographical context(s) dramatically increased

demand for persistent ISR capabilities that LRFW uncrewed aircraft were well positioned to offer while simultaneously and unexpectedly lowering the technological requirements for what constituted a practical LRFW uncrewed aircraft.

As explained earlier in Part III of this dissertation, while the employment of the American AQM-34 (i.e., Ryan Model 147 Lightening Bug) in the Vietnam War defined the realm of possibilities for LRFW uncrewed aircraft technology by the early 1970s, such uncrewed aircraft designs were not very practical in less permissive military-geographical contexts including that which existed in Central Europe. LRFW uncrewed aircraft were for the most part inferior substitutes to existing and prospective crewed aircraft in most roles, including ISR, and could not therefore compete with crewed aircraft for funding and roles. Had the military-geographical contexts in which the leading militaries of the day operated been more permissive, much as Israel's military-geographical context was more permissive relative to that which existed in Central Europe, then the likes of the 1960s vintage turbojet-powered AGM-34, as well as the later propeller-driven Israeli Mastiff and Scout, among other designs, would have likely been more practical and competitive such that these technological developments may have had greater effects on the fielded military capabilities of states.

With the end of the Cold War and the rapid dismantling of what had been large militaries primed for wartime mobilization and high-intensity conventional conflict under the shadow created by large and diverse nuclear arsenals, the bar for a practical LRFW uncrewed aircraft was unexpectedly lowered. Whereas designers and militaries in the 1970s and 1980s had to take into account the survivability, or lack thereof, of uncrewed ISR aircraft in the face of an adversary's combat aircraft and air defenses – to say nothing

of perceived requirements to function in a nuclear battlefield, the end of the Cold War established very different military-geographical contexts that allowed designers and militaries to pursue LRFW uncrewed aircraft for their comparative advantages in the ISR role even though these were no less vulnerable to air defenses. Stated differently, the end of the Cold War amounted to something of a major supply-side shift that lowered the bar for what amounted to a practical uncrewed ISR aircraft capable of competing with crewed ISR aircraft.

To appreciate the pace and magnitude of effects engendered on fielded military capabilities by advances in LRFW uncrewed aircraft technology in the period ca. 1990-2010, it is productive to first summarize some of the changes in military requirements during the 1990s and then examine the implications of advances in key component technologies on the practicality of LRFW uncrewed aircraft and their competitiveness vis-à-vis crewed aircraft. Having done so, we will be well-positioned to examine the implications of the more familiar advances in LRFW uncrewed aircraft technology from ca. 2000 through 2010 on the fielded military capabilities of states.

15.1 Changing Military Requirements in the Post-Cold War Era

While the reunification of Germany dramatically altered the scenarios for which leading militaries primarily prepared, Iraq's invasion of Kuwait and the subsequent American-led intervention in that conflict abruptly created a situation in which the experiences of the Gulf War reshaped military requirements before either the Warsaw Pact or the Soviet Union formally dissolved.

Although aerial bombing during the Gulf War was primarily conducted by crewed combat aircraft employing unguided munitions,[83] uncrewed aircraft technology, broadly defined, played a notable and outsized role in the conflict. Much as Israel had in 1973 and 1982, the United States deployed LRFW uncrewed target aircraft as decoys at the outset of air combat operations to lure Iraqi air defenses into activation so that these could, in turn, be targeted by crewed combat aircraft using anti-radiation missiles such as the AGM-88 HARM. In addition to several dozen ground-lunched LRFW uncrewed target aircraft, the United States employed over a hundred of the smaller ADM-141 TALD air-launched unpowered (glide) decoys for the same purpose, thereby making the Gulf War the first conflict in which large numbers of uncrewed aircraft were employed as decoys to augment crewed combat aircraft in their operations and essentially substitute for crewed combat aircraft as bait.

The Gulf War was also the first conflict since the Second World War in which cruise missiles, of a much more sophisticated and automated variety, were used in large numbers. The United States employed over 300 cruise missiles, primarily ship- and submarine-launched RGM/UGM-109 Tomahawk land-attack cruise missiles, against targets in Iraq.[84] While the number of cruise missiles employed was a rounding error when compared with the total number of munitions of all types employed by American crewed combat aircraft during the conflict, what is notable is that Tomahawk cruise missiles were exclusively launched in the first three days of Operation Desert Storm, which had a duration of forty-three days. As a result, these cruise missiles were employed

[83] Thomas A. Keaney and Eliot A. Cohen, *Gulf War Air Power Survey: Summary Report* (Washington, DC: U.S. Government Printing Office, 1993)

[84] "Air Decoys Fooled Iraq, Navy Reveals," *Chicago Tribune*, September 19, 1991, https://www.chicagotribune.com/1991/09/19/air-decoys-fooled-iraq-navy-reveals/

less to complement crewed combat aircraft but to substitute for most types of crewed combat aircraft – with the exception of the low radar signature (i.e., "stealth") F-117 Nighthawk – in high-risk missions in the opening days of the operation during which Iraq's air defenses and combat aircraft had not been neutralized.

The successful employment of uncrewed aircraft in the form of decoys and cruise missiles by the United States demonstrated that advances in uncrewed aircraft technology, above all in the area of aircraft automation, allowed militaries to both practically complement and substitute for crewed combat aircraft with uncrewed aircraft in a small subset of combat roles. Uncrewed decoy aircraft and cruise missiles, while large fixed-wing uncrewed aircraft, were not, however, reusable, and far fewer large reusable fixed-wing uncrewed aircraft, miniaturized or otherwise, were employed in the conflict. The American RQ-2, an approximately 200-kilogram propeller-driven design of Israeli origin, was successfully employed on a small scale as an uncrewed artillery observer for American Iowa-class battleships operating off the Kuwaiti coastline, and a small detachment was also employed in support of American ground forces operating in the desert.[85] The United Kingdom, for its part, employed a comparatively rudimentary Cold War-era design, the CL-90 (known as the Midge in British service), in small numbers to support artillery units. Beyond these small-scale uses on the part of the United States and the United Kingdom, there is exceedingly little evidence of miniaturized LRFW ISR being employed in the Gulf War, and battlefield ISR missions were almost entirely undertaken by crewed fixed- and rotary-wing aircraft. This small scale of employment is reflective of the reality that small numbers of miniaturized LRFW

[85] Norman Polmar, "Historic Aircraft – The Pioneering Pioneer," *Naval History Magazine* 27, No. 5 (2013), https://www.usni.org/magazines/naval-history-magazine/2013/september/historic-aircraft-pioneering-pioneer

uncrewed ISR aircraft existed anywhere in the world outside of Israel at this time and late 1980s developmental projects had not yet resulted in service-ready uncrewed ISR aircraft for other countries. The less permissive military-geographical context found in Central Europe during the Cold War had, after all, undermined the competitiveness of LRFW uncrewed ISR aircraft relative to crewed ISR aircraft, and this slowed the adoption of technological advances that may have otherwise diffused more rapidly and more widely.

Whatever the role of LRFW uncrewed aircraft in the conflict, Operation Desert Storm, in which American military forces were both overwhelmingly preponderant and played a critical role in degrading Iraqi military capabilities through air and missile strikes, had a dramatic effect on perceptions of military requirements for the post-Cold War era. The war revealed relatively few hitherto "secret capabilities" – most of the equipment and technologies used were known quantities for the better part of a decade if not longer, but many technological advances made by the United States and its allies during the 1970s and 1980s were largely untested in high intensity and sustained combat operations. The conflict therefore served as a proving ground and provided evidence of the practicalities and potential effectiveness of new technologies including uncrewed aircraft of all types from guided air-to-ground munitions to cruise missiles to uncrewed target aircraft and, indirectly, LRFW uncrewed ISR aircraft.

Although the vast majority of air-to-ground munitions employed in the Gulf War were unguided, guided munitions received disproportionate attention from military and public audiences alike.[86] Sustained advances in this supposedly "new" decades-old area of military technology – the United States pioneered the employment of laser-guided

[86] See, for example, Daryl G. Press, "The Myth of Air Power in the Persian Gulf War and the Future of Warfare," *International Security* 26, No. 2 (2001): 5–44.

bombs in the context of the Vietnam War[87] and the first guided bombs were, as previously explained, used in the Second World War – appeared to offer militaries a means of rapidly suppressing air defenses and destroying enemy ground forces without the cost, human and financial, and complexity of major ground operations. Given the backdrop of declining defense expenditures that remained difficult to justify in a context of greatly diminished threat perceptions, leading militaries turned to guided munitions to offset the effects of reductions in force structure.

While leading militaries made large-scale multi-year investments to further develop guided air-to-ground munitions technology, procure such munitions, and equip crewed combat aircraft with the avionics required to effectively employ such munitions, the existence of deployed guided munitions meant little if militaries lacked adequate ISR capabilities to detect and track potential targets in the first place. This dynamic was recognized in the 1980s in the context of ground combat scenarios in Central Europe and spurred a range of developmental efforts toward improving ISR capabilities. This included equipping ground vehicles and aircraft with the latest sensors, the pursuit of more capable crewed and LRFW uncrewed aircraft for the battlefield ISR role (which were not deployed in sizeable numbers by the end of the Cold War), as well as more

[87] John T. Correll, "The Emergence of Smart Bombs," Air Force Magazine (March 2010), https://www.airandspaceforces.com/PDF/MagazineArchive/Documents/2010/March%202010/0310bombs.pdf ; Peter deLeon, *The Laser-Guided Bomb: Case History of a Development* (Santa Monica, CA: RAND Corporation, 1974) ; John Mueller, "The Perfect Enemy: Assessing the Gulf War," *Security Studies* 5, No. 1 (1995): 77-117. ; Stephen Biddle, "Victory Misunderstood: What the Gulf War Tells Us about the Future of Conflict," *International Securit*y 21, No. 2 (1996):139-179. ; Stephen Biddle, "The Gulf War Debate Redux: Why Skill and Technology are the Right Answer," *International Security* 22, No. 2 (1997): 163-174. ; Daryl G. Press, "Lessons from Ground Combat in the Gulf: The Impact of Training and Technology," *International Security*, 22, No. 2 (1997), 137-146. ; Eliot Cohen, "A Revolution in Warfare," *Foreign Affairs* 75, No. 2 (March/April 1996): 37-54. ; Barry D. Watts, *Six Decades of Guided Munitions and Battle Networks: Progress and Prospects* (Washington, D.C.: Center for Strategic and Budgetary Assessments, 2007)

ambitious efforts on the part of the United States that, among other things, led to the development of the E-8 JSTARS, a Boeing 707 airframe equipped with a large synthetic aperture radar (SAR) capable of detecting stationary and moving ground targets located several hundred kilometers away in all weather conditions (i.e., a major gain in combat capabilities that, given the latest in then-current electronics technology, required a very large crewed aircraft to serve as the host platform).

While the Gulf War put into motion several trajectories for capability and technology development, it was the conflicts that followed the dissolution of Yugoslavia that created and sustained a surge in demand for the sustained employment of uncrewed aircraft in ISR roles.[88] The intervention of American and NATO forces in the conflicts transpiring in the Balkans dramatically increased battlefield ISR requirements. Unlike most previous conflicts, including the Gulf War, the countries intervening in the Balkans did so under (politically and diplomatically driven) strict rules of engagement that required military personnel to distinguish between combatants and civilians as well as different groupings of combatants. Unlike the Gulf War, in which the United States-led coalition fought uniformed Iraqi combatants that were primarily operating alongside visually distinct heavy military equipment outside of urban areas, the militaries intervening in the Balkans had the difficult task of distinguishing between both regular and irregular forces of all stripes in conflicts that had many of the hallmarks of civil wars (largely) devoid of large pitched battles between field formations. With no aerial sensor in existence allowing for the discernment of the political allegiances and ethnic identities of distant individuals, the battlefield ISR capabilities developed during the Cold War and

[88] Elizabeth Becker, "Crisis in the Balkans: The Drones; They're Unmanned, They Fly Low, and They Get the Picture," *New York Times*, June 3, 1999, https://www.nytimes.com/1999/06/03/world/crisis-balkans-drones-they-re-unmanned-they-fly-low-they-get-picture.html

employed during the Gulf War, which were oriented toward locating and tracking large mechanized formations, left much to be desired. The conflicts in the Balkans required sustained and ideally persistent ISR operations over combat areas instead of more conventional brief (i.e., snapshot view) peacetime ISR operations over friendly airspace prior to the start of hostilities.

Compounding matters for the intervening militaries, not only did a (brief) overflight by a reconnaissance version of a crewed fighter aircraft provide little in the way of useful information about events that had transpired on the ground (i.e., several hours before captured imagery was processed and analyzed), but political and diplomatic concerns created an environment in which the loss or capture of a crewed aircraft and its onboard human operators was considered unacceptable. While the former Yugoslavia had by no means an advanced military by late Cold War standards (when compared with leading NATO and Warsaw Pact forces), the country operated large numbers of surface-to-air missile systems and air defense artillery pieces. Serbian forces, against whom American and NATO forces primarily operated, controlled many of these air defense capabilities and could find cover and concealment in both urban areas and the rugged and densely vegetated terrain found in the Balkans.[89]

Even though American and NATO forces were, for the most part, not involved in (all-out) combat operations against Serbian forces with the primary exceptions of Operation Deny Flight from 1993 to 1995, Operation Deliberate Force in 1995, and Operation Allied Force in 1999, foreign military aircraft operated in an environment in

[89] This dynamic received considerable attention. See, for example, Barry R. Posen, "The War for Kosovo: Serbia's Political-Military Strategy," *International Security* 24, No. 4 (2000), pp. 39–84.

which aerial operations were largely uncontested at higher altitudes but air defenses, from medium-range surface-to-air missile systems to shoulder-fired anti-aircraft missiles and anti-aircraft artillery, could go undetected and be activated at any time. As a result, American and NATO aircraft on patrol would typically operate at altitudes above 15,000 feet, the approximate ceiling of anti-aircraft artillery fire and shoulder-fired anti-aircraft missiles, and largely operate unhindered as Serbian surface-to-air missile systems infrequently activated their radars and increasingly served as a force-in-being.[90]

In this context, the United States and large NATO militaries turned to LRFW uncrewed aircraft for ISR missions over the Balkans. Whereas crewed combat aircraft could undertake evasive maneuvers, deploy countermeasures, and were equipped with ejection seats, other fixed-wing crewed ISR aircraft (i.e., those that were not versions of combat aircraft) and all rotary-wing ISR aircraft were far more vulnerable to Serbian air defenses and were not designed to operate so far ahead of the frontlines established by friendly ground forces. These unique circumstances created an opening for the employment of LRFW uncrewed ISR aircraft and casualty-averse American and NATO forces deployed available systems in the Balkans. This included late Cold War designs like the turbojet-powered CL-289 (deployed by French and German forces), the propeller-driven BAE Phoenix (deployed by British forces), and the propeller-driven RQ-2 and RQ-5 (deployed by American forces). Notably, the United States also deployed the General Atomics Gnat 750, a derivative of the late 1980s Amber developmental LRFW

[90] While Serbian forces engaged and shot down American and NATO aircraft, the vast majority of crewed aircraft sorties were uneventful. On the paradox of American and NATO dominance at higher altitude and contested dominance at lower altitudes, Barry R. Posen, "Command of the Commons: The Military Foundation of U.S. Hegemony," *International Security* 28, No. 1 (2003): 5–46.; Benjamin S. Lambeth, "Kosovo and the Continuing SEAD Challenge," *Aerospace Power Journal* 16, No. 2 (2002): 8-22.

uncrewed aircraft, for ISR missions over Bosnia from 1994 and later deployed the Gnat 750's larger and more capable derivative and successor, the RQ-1 Predator.[91] These new, significantly larger LRFW uncrewed ISR aircraft were capable of long-endurance missions and could carry a larger sensor payload to a higher altitude and played an increasingly important role in the Balkans in the late 1990s.

While crewed combat aircraft deployed by American and NATO forces far outnumbered LRFW uncrewed aircraft by the time of their intervention in the Kosovo War in 1999, it is instructive that American and NATO forces lost fifteen LRFW uncrewed aircraft and two crewed aircraft, including at least one RQ-1 Predator, during Operation Allied Force, which lasted seventy-nine days.[92] This disparity in aircraft losses can be explained by both the deliberate employment of the relatively small numbers of LRFW uncrewed ISR aircraft in existence on higher-risk missions and the fact that these LRFW uncrewed ISR aircraft were far more vulnerable to Serbian air defenses. Aircraft losses are, however, a poor metric for discerning the relative importance of one type of aircraft or another. The vast majority of sorties in the Kosovo War were carried out by crewed aircraft, and with the exception of cruise missiles and air-launched guided munitions, uncrewed aircraft were not employed in any combat missions.[93]

---

[91] Brian W. Everstine, "Elegy for the Predator," Air & Space Forces Magazine, May 29, 2018, https://www.airandspaceforces.com/article/elegy-for-the-predator/ ; "General Atomics RQ/MQ-1 Predator," Designation Systems, No Date, https://www.designation-systems.net/dusrm/app2/q-1.html

[92] Daniel L. Haulman, "Manned Aircraft Losses Over the Former Yugoslavia, 1994-1999" (Air Force Historical Research Agency, 2015), https://www.afhra.af.mil/Portals/16/documents/Airmen-at-War/Haulman-MannedAircraftLossesYugoslavia1994-1999.pdf?ver=2016-08-22-131404-383

[93] Scholars gave considerable attention to the Kosovo air war. See, for example, Daniel L. Byman and Matthew C. Waxman, "Kosovo and the Great Air Power Debate," *International Security* 24, Bo. 4 (2000): 5–38.; Benjamin S. Lambeth, *NATO's Air War for Kosovo: A Strategic and Operational Assessment* (Santa Monica, CA.: RAND, 2001)

By the turn of the century, uncrewed aircraft, broadly defined, played a small but noticeable and increasingly prominent role in the (new) types of aerial operations undertaken by the United States and leading NATO militaries. After the Gulf War, the United States repeatedly turned to cruise missiles to substitute for crewed combat aircraft in undertaking high-risk (and relatively low-cost) air strikes against targets in Iraq and the former Yugoslavia (as well as targets in Sudan and Afghanistan) so as to eliminate the potential for the death or capture of a human pilot, and the small but growing number of LRFW uncrewed ISR aircraft in existence, including the RQ-1 Predator, were similarly used to conduct both persistent ISR missions and high-risk ISR missions over Iraq and the Balkans.[94] As a result, after decades of technological development, the interplay of (unanticipated) demand-side and supply-side dynamics resulted in an altered military-geographical context in which leading militaries found it practical to pursue LRFW uncrewed aircraft for ISR roles despite the enduring vulnerabilities to air defenses and consequent unsuitability for high-intensity operations against better equipped and more closely matched adversaries. Stated differently, the latest LRFW uncrewed aircraft increasingly affected fielded military capabilities by the turn of the century less as a function of technological change and more as a function of the fundamental shift – a downgrade – in military requirements.

[94] Thom Shanker, "Threats and Responses: Combat; Iraqi Planes Shoot Down American Predator Drone," *New York Times*, December 24, 2002, https://www.nytimes.com/2002/12/24/world/threats-and-responses-combat-iraqi-planes-shoot-down-american-predator-drone.html

15.2 Background Technological Developments

The above summary of the employment of LRFW uncrewed aircraft in the 1990s allows us to understand shifting demand-side factors and the implications these had on the perceived practicality of LRFW uncrewed aircraft for ISR roles. To develop a more holistic understanding of the implications of technological change on fielded military capabilities, we need to closely examine the qualities and limitations of LRFW uncrewed aircraft technology in this timeframe. By so doing, we can discern what effects advances in uncrewed aircraft technology had on fielded military capabilities beyond the new and far more permissive military geographical contexts that characterized American and NATO military operations against Iraq (after the Gulf War) and in the Balkans.

As emphasized repeatedly throughout Part Three of this dissertation, an uncrewed aircraft must have a means of independent navigation to be practical for a variety of military roles. As early as the 1910s, gyroscopes allowed for the stabilization of an aircraft, crewed or otherwise, in flight but did not allow for operation without a trained human pilot at the controls. By the 1940s, automation technology had become more advanced, but it took the refinement of inertial navigation systems over the 1950s and 1960s to allow an uncrewed aircraft to independently navigate to and from a predesignated area. By the 1970s, advances in solid-state electronics facilitated improvements in computation, thereby supporting the development of more advanced autopilots, and the coming of full duplex radio datalinks and electro-optical sensors allowed for the practical operation of remotely piloted aircraft at distances up to 150 or so kilometers away.

To independently operate much beyond that range, an uncrewed aircraft had to be fully automated in the manner of a cruise missile such as the American Tomahawk. During the 1970s and 1980s, this entailed the employment of multiple complex apparatus for navigation and guidance, a dynamic that made uncrewed aircraft capable of independent operations at long range very expensive. These late Cold War navigation and guidance systems, which were for the most part restricted to operating at lower altitudes, could not, however, be paired with long-range datalinks capable of operations beyond the line-of-sight of a ground station (i.e., limitations in radio/datalink technology compounded the limitations of aircraft automation technology). As a result, uncrewed aircraft operating at long range and/or at low altitudes could neither share sensor data nor offer a remote human operator – not necessarily a human constantly in the loop – both feedback and the capacity to send new commands to the uncrewed aircraft. By the 1990s and during the two decades that followed, however, sustained technological advances in five areas made the operation of LRFW uncrewed aircraft over long distances increasingly practical.

First, automation technology progressively improved such that flight controls could be increasingly automated. Perhaps the most important example of this development in a deployed aircraft was the American RQ-4 Global Hawk, an uncrewed high-altitude ISR aircraft (in the vein of a U-2) capable of alternating between fully automated operation (i.e., following preprogrammed flight paths) and real-time human control via a radio datalink. Just as with cruise missiles, fully automated operation (i.e., no human in the loop) required expensive onboard computers and navigation systems (of a different variety than those used on cruise missiles) and the remotely piloted mode of

operation (i.e., human in the loop) remained more affordable and more practical for most applications. As a result, remote piloting through a radio datalink, supported by some level of onboard automation of flight controls, was by far the most common mode of operation throughout the ca. 1990 to 2010 timeframe (the RQ-4 was an exquisite design not reflective of essentially any other LRFW uncrewed aircraft even as it defined the technological frontier).

Over this timeframe, progressive advances in aircraft automation technology allowed for moderate levels of onboard automation that could at least function as a backup in the event that the real-time control radio link was disrupted or lost. In some LRFW uncrewed aircraft designs, onboard automation allowed the aircraft to enter a holding pattern (until it ran out of fuel). In other, more advanced designs, the aircraft could fly back toward the launch area, an action that improved the prospect of restoring the radio link. Without these advances in aircraft automation technology, disruptions to radio links would result in uncrewed aircraft flying without a human at the controls until the aircraft lost altitude and impacted terrain. Given the cost of the airframe and onboard sensors and datalinks, advances in automation over the ca. 1990-2010 timeframe made LRFW uncrewed ISR aircraft increasingly reliable and therefore increasingly practical and competitive relative to crewed ISR aircraft.

Second, advances in radio technology facilitated the development of less expensive and higher bandwidth radio datalinks and this is intimately tied to a third area of technological advancement, that of sensor technology. Advances in radio datalink technology meant that uncrewed aircraft could transmit in real-time data collected by increasingly capable sensors, including electro-optical sensors, and this sensor data

offered a remote human operator greater sensory acuity (i.e., feedback) to navigate the aircraft from afar. With low-resolution real-time imagery from earlier electro-optical sensors, remote pilots had difficulty navigating over long distances and were as such restricted to operating in high-visibility flying conditions. Advances in infrared sensor technology, however, facilitated the operation of remotely piloted aircraft at night and in low visibility conditions. By circa 2010, compact synthetic aperture radar (SAR) became available and offered suitable uncrewed aircraft the means to collect imagery in all weather conditions. Together, advances in radio datalink and sensor technology allowed for increasingly sophisticated and practical LRFW ISR aircraft that could better replicate the capabilities offered by crewed ISR aircraft.

Neither improvements in automation nor datalinks, however, offered an onboard or remote means of navigation, and all aircraft, and indeed all vehicles, whether crewed or uncrewed, require accurate location information to practically operate over long distances and/or long periods of time. In uncrewed aircraft that lack an onboard human operator equipped with a bounded "navigational computer" and eyes through which to make visual references, a means to accurate navigation becomes even more important. A fourth area of technological advancement dramatically increased the practicality of operating uncrewed aircraft over long distances. By the mid-1990s, space-based radio navigation systems, of which the American Global Position System (GPS) is the best-known example, offered global precision navigation and timing.[95] As a result, an aircraft, whether crewed or uncrewed, equipped with a GPS receiver could access accurate location data in real-time and GPS could also be used to complement increasingly

[95] Paul E. Ceruzzi, *GPS* (Cambridge, MA: The MIT Press, 2018)

accurate inertial navigation systems by accounting for accumulating drift in these systems. Together, the combination of GPS and accurate inertial navigation systems facilitated advances in aircraft automation technology as uncrewed aircraft could largely dispense with a remote human pilot for the purposes of navigation to and from a target area.[96]

These four areas of technological development greatly expanded the potential for long-range (and long endurance – endurance being aggregate flight range over time) operations using LRFW uncrewed aircraft, but one problem endured. A fully automated uncrewed aircraft equipped with GPS and an inertial navigation system is desirable in theory, but in practice, its military utility is limited for most roles, other than a single-use strike role in the manner of a cruise missile, if it cannot share sensor data with and receive updated commands from a remote human operator. High bandwidth radio datalinks are line-of-sight systems and even a high-altitude uncrewed aircraft on a long-range mission will eventually operate beyond the line-of-sight of the radio antenna of the ground station to which it transmits sensor data and from which it receives new commands from a remote human operator. A fifth area of technological development offered a means of untethering uncrewed aircraft from their ground stations. By the 1990s, communications satellites, both civilian and military, facilitated real-time global radio communications. Over time, the bandwidth of satellite communication systems increased, and this allowed uncrewed aircraft to operate increasingly sophisticated sensors and transmit increasingly large volumes of sensor data in real-time.

[96] James M. Hasik, *The Precision Revolution: GPS and the Future of Aerial Warfare* (Annapolis, MD: Naval Institute Press, 2002)

The combination of advances in automation, radio datalinks, sensors, GPS, and satellite communications meant that by the late 1990s, it became possible for an LRFW uncrewed aircraft to accurately navigate toward a distant target area (up to the maximum range of an airframe), transmit sensor data in real-time beyond the line-of-sight of a ground station by using a satellite as a communications relay, return to a designated area for landing such that the aircraft could be prepared for future missions, all while responding to updated commands from remote human operators throughout the mission. Stated differently, by the late 1990s, background technological advances allowed LRFW uncrewed ISR aircraft to more effectively compete with crewed ISR aircraft for funding and roles.

15.3 New LRFW Uncrewed Aircraft Designs

In this context, both military-geographical and technological, multiple countries, including the United States and Israel, the technology leaders, pursued new, more capable LRFW uncrewed aircraft designs that would in time be exported to and emulated by other countries. Whereas the Scout and Mastiff were relatively small and rather primitive designs, the LRFW uncrewed aircraft that Israel introduced into service during the 1990s were progressively larger and heavier aircraft equipped with more powerful engines. As a result, these propeller-driven aircraft, which had airframes optimized for glide performance rather than maneuverability or speed (which, alongside payload, constitute the performance parameters that combat aircraft designs prioritize), could carry more fuel and/or payload and operate further and/or longer – and typically at higher altitudes – than preceding LRFW uncrewed aircraft. The Israeli IAI Heron, for example, had a maximum

take-off weight of around 1,100 kilograms and could carry a payload of 250 kilograms, which was adequate for a suite of sensors comparable to those used on smaller, lower-end crewed ISR aircraft. The smaller Israeli Elbit Hermes 450 weighed around 450 kg and could carry a payload of around 100 kilograms. The increased weight of these aircraft, which greatly exceeded the approximately 150 to 200 kilograms maximum takeoff weight of 1980s miniaturized LRFW uncrewed ISR aircraft designs, reflected greater fuel capacity and the additional mission systems required for onboard automation, sensors, and datalinks, as well as the larger and more powerful engines required to maintain performance and facilitate operations at higher altitudes.

The United States, for its part, fielded a version of an Israeli design, the IAI Hunter, as the RQ-5, an aircraft falling between the Heron and Hermes 450 in performance, and, after a failed procurement of the American RQ-6 design, later fielded the RQ-7 Shadow, all of these were propeller-driven designs using small piston engines. While the RQ-7 was a smaller aircraft weighing around 150 kilograms with a commensurately smaller payload capacity, the United States also deployed small numbers of the Gnat 750, a derivative of the late 1980s Amber prototype aircraft, in the early 1990s and further developed this design in a developmental effort that resulted in the RQ-1 Predator, an aircraft that entered service in 1995.

Weighing around 950 kilograms, the RQ-1 was in the same weight and performance class as the Israeli Heron, a detail best reflected in the fact that both aircraft were equipped with a single Austrian Rotax 914 piston engine. Generating 115 horsepower, or just over two-thirds of the output of a Cessna 172's piston engine (the Cessna 172 is a four-seat utility aircraft that has dominated general aviation since the late

1950s and is the most produced crewed aircraft in history), the Rotax 914 allowed the RQ-1 Predator to attain a maximum speed of around 220 kilometers per hour, which was about as fast as a Cessna 172 or a UH-1 Iroquois helicopter (i.e., a Huey). What the RQ-1 lacked in speed, however, was outweighed by the combination of a relatively efficient engine and an airframe optimized for glide performance. With its unswept ~15-meter-long high aspect ratio wing, the RQ-1 could operate at around 25,000 feet, around 10,000 feet higher than smaller contemporary miniaturized LRFW uncrewed aircraft, and, when equipped with a satellite communications link and GPS receiver, it could loiter over an area for over 20 hours and transmit sensor data in real-time. As a result, the RQ-1 could effectively compete with the military versions of small propeller-driven aircraft, such as the American RC-12, an ISR version of the Beechcraft Super King Air (a twin-turboprop powered thirteen-seat transport aircraft), for ISR roles in permissive airspace.

While some countries imported American and/or Israeli LRFW uncrewed aircraft for ISR roles from ca. 1990 to the early 2000s, others pursued designs of their own, a process that was hindered by the fact that most leading militaries, with the notable exceptions of the United States and Israel, had dramatically reduced military spending, including expenditures toward research and development following the end of the Cold War. France, for example, developed and fielded the Sagem Sperwer and Germany developed and fielded the Rheinmetall KZO, both of which were smaller designs intended to undertake ISR missions in support of ground formations and both of which were only built in small numbers. By the early 2000s, other countries were developing experimental reusable fixed-wing uncrewed ISR aircraft of their own but with the exception of a handful of countries such as China and Iran, these tended to be smaller

designs often weighing less than one hundred kilograms (i.e., uncrewed aircraft designs that are beyond the scope of this dissertation).

15.4 The Deficiencies of the Latest LRFW Uncrewed ISR Aircraft: The Case of the RQ-1 Predator

Over the 1990s and into the mid-2000s, LRFW uncrewed aircraft technology gradually diffused, and the United States, the leader across a wide range of relevant technologies including, not unimportantly, both satellite communications and GPS, returned to its longstanding practice of undertaking highly ambitious developmental efforts aimed at creating LRFW aircraft that could better compete with crewed aircraft across a wider range of ISR missions. While the RQ-1 Predator was a useful and effective ISR asset, its primary positive attributes were its ability to undertake long-endurance missions and its ability to do so without putting an onboard human operator at risk. In terms of speed, payload, operating altitude, and maneuverability, among other performance parameters, the RQ-1 left much to be desired, and it is important to recognize that the RQ-1 Predator was the most capable LRFW uncrewed ISR aircraft in military service at the turn of the century.

A small piston engine and a slow maximum speed, while contributing to the RQ-1's long endurance, meant that it would take around two and a half hours for an RQ-1 to reach a distant mission area located around five hundred kilometers away. In contrast, a crewed fighter aircraft or a crewed business jet equipped with sensors with equal if not greater performance than those found on the RQ-1 could reach the target area in around half an hour. The RQ-1 was, as such, unsuited to undertaking ISR missions on short

notice and was most effective when it was already flying over the target area. In practice, this meant that LRFW uncrewed aircraft equipped with piston engines were more practical for some ISR roles and less practical in others and militaries could not therefore use uncrewed aircraft like the RQ-1 to fully substitute for crewed ISR aircraft let alone other military aircraft. Inadequate speed was not, however, the only deficiency of the RQ-1 and similar LRFW uncrewed ISR aircraft. A maximum payload of around 200 kilograms allowed an RQ-1 to carry larger, heavier, and more capable sensors than prior miniaturized LRFW uncrewed ISR aircraft but the RQ-1 could not compete with converted small passenger aircraft, larger converted business jets, crewed combat aircraft, or highly specialized ISR aircraft like the U-2, all of which can carry at least 1,000 kilograms of sensors (i.e., the entire weight of the RQ-1), in this important performance parameter.

All things equal, the greater the payload capacity in terms of weight and volume, the greater the quality and/or number of sensors that can be carried onboard an aircraft, whether crewed or uncrewed. With a payload of just 200 kilograms, the RQ-1 was primarily limited to a single sensor ball featuring a visible light electro-optical camera and an infrared camera for operations in low light conditions (as well as a laser designator in later versions). While this was a useful sensor payload, a single sensor (ball) is limited in terms of field of view, and zooming toward a distant area further restricts the field of view, a dynamic best characterized as looking through a straw that progressively decreases in diameter as the view is magnified. In a larger aircraft with greater payload capacity, multiple and/or more capable sensors could be carried. Compounding matters, the RQ-1, equipped as it was with a modest piston engine, could not operate above

25,000 feet. Essentially all electromagnetic sensors, whether active or passive and irrespective of the frequency of operation, are line-of-sight systems, and the service ceiling of the RQ-1 restricted its field of view such that it could not compete with larger, heavier aircraft equipped with more powerful turboprop or turbofan engines and comparable let alone superior sensors.

Whatever the RQ-1's qualities as an ISR aircraft, its design did not lend itself to practical operations outside of permissive airspace. The high aspect ratio wings used on the RQ-1 and other miniaturized LRFW uncrewed ISR aircraft offered excellent glide performance and, when paired with an efficient but modest engine with limited reserve power at cruise speed, offered the capacity to undertake long endurance missions, but the trade-off was that the RQ-1 was not a maneuverable aircraft. As a result, it was poorly positioned to undertake evasive maneuvers. With anti-aircraft missiles flying at supersonic speed (typically Mach 2 to Mach 4) and capable of undertaking high-G maneuvers far beyond human physiological limits and, as such, the (mutable) structural limits of crewed combat aircraft, uncrewed aircraft in the vein of the RQ-1 were extremely vulnerable to adversary air defenses and combat aircraft and this vulnerability only increased with operating altitude which was, in effect, correlated with the RQ-1's utility as an ISR aircraft. If the RQ-1 had a far greater payload capacity, it could have been equipped with countermeasures such as flares (to counter missiles with infrared guidance) and chaff (to counter radar-guided missiles) as well as a wide range of electronic countermeasures from radar warning receivers and missile approach warning systems to jamming pods and towed decoys. Such systems were standard on crewed combat aircraft and high-end crewed ISR aircraft, but the RQ-1 lacked the spare volume

and payload capacity – and the spare power and cooling capacity – to be equipped with such systems. Without onboard countermeasures and with no capacity to undertake meaningful evasive maneuvers, the RQ-1 was practical only in a permissive operating environment and would otherwise be a very expensive short service life ISR aircraft if employed against more capable adversaries.

To address some of these limitations, the United States pursued the development of a larger derivative of the RQ-1 that was equipped with a more powerful turboprop engine (i.e., a gas turbine engine). The resulting aircraft, initially called Predator B and later the MQ-9 Reaper, first flew in February 2001. Whereas the Rotax 914 piston engine on the RQ-1 generated 115 horsepower, the MQ-9 was equipped with a 900-horsepower turboprop engine of design widely employed in small transport aircraft such as the nineteen-seat Dornier 228 in a two-engine configuration as well as highly maneuverable military trainer aircraft like the Short Tucano in a single engine configuration. Equipped with this much more powerful engine, the MQ-9 attained a maximum take-off weight exceeding 4,500 kilograms and could carry a combination of up to 2,500 kilograms of fuel or payload (the greater the fuel load carried and the smaller the payload – particularly wing-mounted payload, all things equal, the longer the endurance of the MQ-9). The more powerful turboprop engine also allowed the MQ-9 to reach a maximum speed of around 370 kilometers per hour and a maximum operating altitude of 50,000 feet. To appreciate what this meant in practice and the trade-offs associated with the MQ-9's retention of an airframe optimized for a high glide ratio, it is productive to contrast the MQ-9 with another military aircraft equipped with a similar turboprop engine.

The Short Tucano, a version of the Brazilian Embraer EMB-312 Tucano developed for the United Kingdom, is a highly maneuverable two-seat aircraft used by the British military to train pilots to a level required to advance to training on jet-powered trainer aircraft and, afterward, on fighter aircraft. First flown in 1986, the Short Tucano is relevant to the present purposes because it utilizes a version of the Garret (later Honeywell) TPE331 turboprop engine also used on the MQ-9, albeit a version with a higher maximum power output of 1,150 horsepower. While the Short Tucano is much more maneuverable (being capable of 7 G maneuvers – human physiological limits, not mutable structural limits, restrict crewed fighter aircraft to 9 G maneuvers), slightly faster, and substantially lighter (with a maximum takeoff weight of under 3,300 kilograms), it lacks the high aspect ratio wings required to operate for as long or as high as the MQ-9. As a result, the Short Tucano can operate for up to four hours or so – well within human physiological limits – and has a service ceiling of 34,000 feet. If the United States or any other country was interested in an LRFW uncrewed aircraft with performance comparable to the Short Tucano, one more comparable to crewed combat aircraft (which are typically powered by jet engines), it could have developed such an aircraft or possibly converted existing crewed aircraft including the Short Tucano to an uncrewed configuration. Such an uncrewed aircraft would, however, be poorly optimized for ISR roles and would be less competitive as a substitute for crewed ISR aircraft than the uncrewed MQ-9 Reaper.

This brief discussion highlights that although technological advances allowed for the development of miniaturized LRFW uncrewed aircraft design from the late 1970s onward, the pursuit of greater payload, range, and operating altitude – performance

characteristics required to better compete with crewed ISR aircraft for funding and roles – meant that uncrewed aircraft with the performance of the MQ-9 or greater would be as nearly as large and heavy as comparable crewed military aircraft using the same propulsion systems and were as such no longer miniaturized or relatively inexpensive aircraft designs. Stated differently in the manner of a "citrus-to-citrus" comparison, the primary distinction between the likes of an MQ-9 and the likes of a Short Tucano, which essentially shared an engine, beyond the absence of an onboard human operator in the former, was that the MQ-9 had an airframe characterized by high aspect ratio glider-like wings that were optimized for glide performance (endurance) and service ceiling at the cost of other flight characteristics and performance parameters. The capacity to undertake long endurance missions was, in other words, neither an intrinsic attribute of not having a human pilot on board nor was it without cost or trade-offs, and this affected the implications of advances in LRFW uncrewed aircraft technology in less permissive military-geographical contexts.

15.5 The Pursuit of More Capable LRFW Uncrewed ISR Aircraft

Impressive though the MQ-9 Reaper was as an ISR aircraft and notwithstanding its improvements over the RQ-1/MQ-1 Predator, it was primarily competitive with crewed ISR aircraft in terms of maximum operating altitude,[97] endurance (which, as explained, comes at the cost of major design trade-offs), and, of course, the fact that its operation did not entail putting the life of any onboard human operator(s) at risk. In terms of essentially

[97] Uncrewed propeller-driven aircraft can operate at higher altitudes than crewed propeller-driven aircraft using the same engines because uncrewed aircraft do not have to maintain cabin pressurization (i.e., uncrewed aircraft are not subject to human physiological limitations).

every other performance parameter, even the more capable MQ-9 was inferior to existing and prospective crewed aircraft, and this even applied to operating altitude as the crewed U-2 could operate at over 70,000 feet and do so while carrying a larger, heavier, and incomparably more sophisticated payload than the MQ-9. As part of its long-standing – decades-long – ambition of developing an LRFW uncrewed aircraft that could match or exceed the performance of crewed aircraft in ISR roles, the United States military, buoyed as it was with access to unmatched resources, financial and otherwise, undertook two highly ambitious developmental efforts that resulted in prototypes successfully completing their respective first flights by 1998, which is to say over a year before the start of Operation Allied Force, the NATO intervention in the Kosovo War, and over three years before the 2001 al-Qaeda attacks in the United States and subsequent invasion and occupation of Afghanistan.

The first and more ambitious effort resulted in the RQ-3 Dark Star, one of the first LRFW uncrewed aircraft to have an airframe optimized for a low radar signature, which was developed to meet a long-standing (i.e., Cold War era) American requirement for an LRFW uncrewed aircraft capable of undertaking high-risk ISR missions in contested airspace.[98] Progressive advances in combat aircraft and air defense technology, as well as political and diplomatic concerns, led the United States to largely abandon this ISR role during the Cold War following the shooting down of a U-2 over the Soviet Union in 1960. By the late 1980s, however, the United States pioneered low radar signature (i.e., "stealth") aircraft technology, and this incentivized a (reportedly unsuccessful)[99]

[98] Ehrhard, *Air Force UAVs*.
[99] For a useful summary, see Rebecca Grant, *The Radar Game: Understanding Stealth and Aircraft Survivability* (Arlington, VA: Mitchell Institute for Aerospace Studies, 2010)

ambitious development effort aimed at producing a low radar signature LRFW aircraft that could, in wartime, undertake ISR missions over the Soviet Union (reportedly to target Soviet nuclear-armed mobile missile launchers). What became the RQ-3 in the 1990s was reportedly a less ambitious successor for this little-known late Cold War developmental effort.[100] Weighing over 3,800 kilograms, the RQ-3 was not a small aircraft, and its turbofan engine allowed it to operate at up to 45,000 feet. What distinguished the RQ-3 from prior LRFW uncrewed aircraft except the Boeing Condor was that the onboard automation technology was reportedly capable of independently undertaking preprogrammed missions without a human in the loop. Although several prototypes of the RQ-3 were built and tested, the development effort encountered budgetary and technical challenges and was canceled in 1999 such that the development of the RQ-3 expanded the realm of possibility by extending the technological frontier without engendering any discernable shift in fielded military capabilities.

While the RQ-3 was designed to operate in military-geographical contexts featuring advanced air defenses, the concurrently developed RQ-4 Global Hawk was designed to undertake ISR missions in permissive airspace. Leaving aside the late 1980s Boeing Condor, which was a technology demonstrator, the turbofan-powered RQ-4 constituted the archetype for a high-altitude long-endurance LRFW uncrewed ISR aircraft. Like the crewed U-2, the RQ-4 was designed with long high aspect ratio wings – a wingspan of 35 meters – to optimize endurance when operating at an altitude of up to 60,000 feet. Unlike the RQ-3, the RQ-4 had an otherwise conventional design that was not optimized for a low radar signature. Whereas the RQ-3 never entered service, the RQ-

[100] Ehrhard, *Air Force UAVs*.

4, itself a highly automated aircraft capable of operating without a human in the loop, was pressed into service in 2001 until available prototype aircraft could be replaced with later, more refined production standard versions of the design.

With the RQ-4, the United States realized its decades-long objective of developing an LRFW uncrewed aircraft that could take on some of the roles of the crewed U-2 and do so with an aircraft that did not put a human pilot at risk, and which could remain airborne for longer than human physiological limitations allowed. The RQ-4 was, in other words, the successor to the ambitions reflected in the development of YQM-94 and YQM-98 from the 1970s Compass Cope program and Boeing Condor, which was developed in the 1980s. Although the RQ-4 entered service, the aircraft had significant limitations and a complicated development and service history characterized by a high loss rate of early production aircraft, a dynamic that reflected upon the immaturity of the onboard automation technology at the time it was pressed into service in 2001. Moreover, the RQ-4, a large aircraft approaching the size and weight of the crewed U-2 – around 12,000 kilograms and 18,000 kilograms, respectively – with a price tag to match, would never supplant the U-2 in American service.[101]

15.6 The Plateauing of LRFW Uncrewed Aircraft Technology

While advances in LRFW uncrewed aircraft technology in the ca. 1990-2010 timeframe improved the ISR capabilities of states and thereby engendered non-zero shifts in fielded military capabilities, a holistic understanding of the magnitude of engendered gains in

[101] Valerie Insinna, "Air Force's RQ-4 Global Hawk drones headed for retirement in FY27," *Breaking Defense*, July 27, 2022, https://breakingdefense.com/2022/07/exclusive-air-forces-rq-4-global-hawk-drones-headed-for-the-boneyard-in-fy27/

fielded military capabilities requires consideration of two important and interrelated dynamics that suggest that the capability gains engendered by technological advances made from the 1990s onward were more limited than widely thought.[102]

First, LRFW uncrewed ISR aircraft were almost exclusively employed in military-geographical contexts characterized by a low or non-existent threat from adversary air defenses in the ca. 1990-2010 timeframe. This poses a major analytical challenge when attempting to discern the implications of advances in LRFW uncrewed aircraft technology on the fielded military capabilities of states. As explained repeatedly throughout this dissertation, the military effects of technological change are coproduced by the actions taken and forgone by adversaries, and the more permissive the military-geographical context, the greater the effects of technological change are likely to be, all things equal. As such, we should be very cautious when attempting to extrapolate generalizable inferences from the empirical record of the employment of LRFW uncrewed aircraft in environments characterized by a low or non-existent threat from adversary air defenses such as the American-led wars in Afghanistan and (post-invasion) Iraq. For this reason, this chapter has primarily focused on the qualities and limitations of the latest advances in LRFW uncrewed aircraft technology and how such aircraft compared with crewed aircraft. By so doing, this chapter highlights the enduring limitations of LRFW uncrewed ISR aircraft technology as of 2010.

Second, while technological advances in the areas of automation, sensors, and datalinks meant that LRFW became increasingly sophisticated and increasingly capable

[102] A productive contrast can be drawn between this characterization of the military effects of technological change in this area and the tenor of alarmist writings that dominated academic discourse on the implications, broadly construed, over the 2010s.

in the undertaking of ISR roles over the ca. 1990-2010 timeframe, LRFW uncrewed aircraft technology nevertheless largely *plateaued*, a dynamic that has received surprisingly little attention amidst the growing chorus of writings focusing on the purportedly "revolutionary" effects of advances in uncrewed aircraft technology. To appreciate this dynamic of plateauing, it is important to focus on the airframes of LRFW uncrewed aircraft in this timeframe rather than the onboard automation systems, sensors, and datalinks, essentially all of which can be readily upgraded in existing uncrewed aircraft designs on the basis of background technological developments, and which can, for the most part, also be applied to both existing and new crewed aircraft designs.

As explained previously in the context of multiple LRFW uncrewed designs and most extensively in the context of the archetypical American RQ-1 Predator and MQ-9 Reaper aircraft, the physical characteristics, and specifications of all aircraft, whether crewed or uncrewed, determine the roles for which aircraft are both practical and competitive. In the ca. 2000-2010 timeframe, the vast majority of LRFW uncrewed aircraft in existence were either designs dating from the 1990s – the likes of the Israeli Searcher, Hermes 450, and Heron and the American RQ-7, RQ-1/MQ-1, and MQ-9 – or close analogs of these aircraft. There were, in other words, very few substantial advances in the area of airframes and aircraft performance in LRFW uncrewed aircraft technology during this timeframe and this limited the roles for which LRFW uncrewed aircraft were both suitable and competitive.

While wartime necessity may have driven the likes of the United States to prioritize the production of existing designs over the development of new uncrewed aircraft – a dynamic for which there is scant evidence given the profusion of research and

development efforts undertaken by the United States in this period that had no bearing whatsoever on combat operations in Afghanistan, Iraq, and elsewhere – I contend that LRFW uncrewed aircraft technology plateaued as it encountered technological constraints imposed by the enduring limitations of automation technology. These limitations in automation technology meant that (A) new clean sheet designs tended not to offer substantial advantages over existing LRFW uncrewed aircraft designs upgraded to incorporate the latest in sensors and datalinks and (B) LRFW uncrewed aircraft remained primarily restricted to ISR roles or, at most, armed ISR roles, while operating in permissive airspace because limitations in onboard automation made the undertaking of other, more demanding roles infeasible. As a result, there was little scope for improving the performance of LRFW uncrewed aircraft by pursuing new airframes optimized for roles beyond ISR.

To understand this argument about the plateauing of LRFW uncrewed aircraft technology, we must return to the fundamentals. As explained earlier, all things equal, an uncrewed aircraft has only two *intrinsic* advantages over a crewed aircraft with otherwise equal performance specifications (speed, payload, range, rate of climb, roll rate, etc.) and is equipped with identical sensors and datalinks. First, the absence of an onboard human operator means that there are no human physiological limits on the endurance of the aircraft, fuel permitting, or the maneuverability of the aircraft, airframe permitting.[103] Second, the absence of an onboard human operator means that there is no potential for the loss of a human operator. Beyond these qualities and areas of comparative advantage, an uncrewed aircraft is just another type of aircraft for which performance parameters

[103] This has long been an issue in aviation. Timothy P. Schultz, *The Problem with Pilots: How Physicians, Engineers, and Airpower Enthusiasts Redefined Flight* (Baltimore, MD: Johns Hopkins University Press, 2018)

determine the suitability for various roles. To operate as effectively as a crewed aircraft, an uncrewed aircraft must, however, have onboard automation systems – or at least a datalink that facilitates real-time remote piloting – that can fully substitute for an onboard human pilot. If aircraft automation technology is unable to fully substitute for an onboard human pilot, then the uncrewed aircraft will face restrictions to its mode of operation independent of what the airframe is capable of. As a result, LRFW uncrewed aircraft designs have airframes that match the capacity of their onboard automation and vice versa. This dynamic explains why, as of 2010, LRFW uncrewed aircraft were optimized for ISR roles and new clean sheet designs did not offer substantial advantages over existing LRFW uncrewed ISR aircraft designs, thereby engendering the dynamic of plateauing.

To better understand the implications of plateauing in practice, it is productive to turn to the RQ-1/MQ-1 Predator and its mode of operation. During a typical ISR mission, an aircraft, whether crewed or uncrewed, tends to maintain level flight so as to not interfere with the operation of the onboard sensors. During the Cold War, the likes of the American RF-4, a reconnaissance version of a fighter aircraft with built-in cameras that primarily captured overhead – as opposed to oblique – imagery, would fly toward and over a target area, capture imagery, and then return to land. If, for some reason, a second pass was required, the onboard pilot could navigate the aircraft, fly back toward the target area, and undertake another level flight to capture overhead imagery. When undertaking such operations in peacetime, the ISR mission was for the most part a navigational exercise. In wartime, however, the onboard pilot would need to rapidly respond to a fast-changing environment by maneuvering to avoid radar coverage once alerted by the

onboard radar warning receiver, altering the flight path in response to newly detected air defenses, and navigating in response to terrain and weather encountered as a result of its forced rerouting. In crewed aircraft with sensors not restricted to detecting objects underneath the aircraft, such as a radar-carrying E-3 AWACS aircraft, which typically operate at higher altitudes, the onboard pilot had the option of loitering over the area and would typically maneuver the aircraft on a figure-eight or racecourse-shaped holding pattern. In peacetime, such holding patterns were routine and little more than a navigational exercise but in wartime, the pilots of aircraft like the E-3 had to rapidly adapt to fast-changing circumstances to avoid being shot down by fast enemy aircraft equipped with long-range air-to-air missiles. The likes of the RF-4 and E-3 ISR aircraft could *adapt* in such a manner because the onboard human pilot(s) had a capacity, however bounded, to quickly respond to external stimuli and changing opportunities and threats.

To credibly supplant crewed aircraft, uncrewed aircraft required onboard automation systems that could fully substitute for an onboard human pilot. Given technological limitations in terms of onboard automation technology and, no less importantly, communications technology, this was extremely challenging and forced designers and militaries to pursue uncrewed aircraft for a subset of ISR roles characterized by limited maneuvering, operations at relatively slow speeds, and operations at higher altitudes. To appreciate this dynamic, we must reflect on how a remotely piloted aircraft is operated. While radio signals are transmitted – sent and received – at the speed of light, even milliseconds of latency accumulate in significant effects given the following sequence of events: an onboard camera detects an object, a

radio datalink transmits footage of an object to a remote human pilot, the human pilot discerns this information and processes it, the human pilot makes control inputs in response, the control inputs are transmitted over radio to the uncrewed aircraft, the aircraft responds to the control inputs, and the cycle repeats itself for essentially every second of the flight unless the aircraft is placed in a (basic) supervised autopilot mode or, ideally, if the aircraft can independently discern and respond to changing external stimuli without constant human supervision and intervention over the controls (i.e., without a human constantly in the loop). Given this constantly looping sequence of events, milliseconds of latency rapidly accumulate and pose a major obstacle to the operation of remotely piloted uncrewed aircraft in complex scenarios and environments.

The greater the control latency, the smaller the margin of error in operating a remotely piloted uncrewed aircraft, and the greater the risk of accidents and loss of control, all things equal. A highly maneuverable uncrewed aircraft, for example, is likely to be subject to even the temporary loss of communication with its remote human pilot not least because maneuvering interferes with the line of sight of an onboard radio used to communicate with a satellite system and/or a distant ground station. The faster the aircraft and the lower the altitude at which it operates, the more severe the possible consequences of such a loss of communications and, as such, a loss of human control (without an onboard automated backstop to take over). A high-performance fighter aircraft like the F-16, for example, can climb at a rate of several tens of thousands of feet per minute and can descend at a rate of around 10,000 feet per minute while undertaking up to 9 G maneuvers. If a remotely piloted LRFW uncrewed aircraft with performance comparable to an F-16 was subject to a disruption of its communications link at lower

altitudes, it could crash before the human operators could regain control over the aircraft as a result of control latency. With a cruise speed of around 1,000 kilometers per hour, the aircraft could displace around 16.7 kilometers per minute or around 280 meters per second. At such speeds, a net control latency – including human decision-making speed and manipulation of the controls – of just 3 seconds would mean that the aforementioned F-16 at cruise speed could horizontally displace over 800 meters or vertically displace 500 or more meters.

When employed in a relatively slow LRFW uncrewed ISR aircraft maintaining a level flight at high altitudes, the effects of the interplay of latency and the limitations of aircraft automation were moderated. If, however, designers pursued an LRFW uncrewed aircraft for roles that required greater maneuvering, not only would the likelihood of the loss of the radio command link – and the loss of the aircraft – increase, but the effects of the interplay of latency and the limitations of aircraft automation would be amplified. By operating at higher altitudes, slower speeds, and without much in the way of maneuvering, designers benefited from a greater margin of safety that allowed for the more practical and more reliable employment of LRFW uncrewed aircraft. For these reasons, "slow and steady" higher-altitude ISR roles were a perfect fit for the limited capacity of aircraft automation technology to fully substitute for an onboard human pilot and vice versa. Absent major advances in aircraft automation technology, LRFW uncrewed aircraft technology plateaued, and this limited the effects of technological change on fielded military capabilities given that LRFW uncrewed aircraft were only practical for ISR (and armed ISR roles in permissive operational contexts) and could not, as such, supplant other types of crewed military aircraft.

While international relations scholars, among others, discussed and debated the implications of LRFW uncrewed aircraft and successive advances in this area of military technology by circa 2010, the observable military effects were far more limited given that LRFW uncrewed aircraft technology was at an impasse. The limitations of onboard automation technology largely restricted LRFW uncrewed aircraft to ISR and armed ISR roles in the narrow context of operations against primarily non-state adversaries with negligible if not non-existent air defense capabilities. For this reason, LRFW uncrewed aircraft could not be practically or productively substitute for or supplant crewed military aircraft beyond the niche roles of persistent ISR and armed ISR in permissive airspace.[104] In this respect, the state of debate on the implications of advances in uncrewed aircraft technology as of ca. 2010 was very much speculative and premature, and advances in LRFW uncrewed aircraft technology over the ca. 1990-2010 timeframe did not engender sizeable and abrupt shifts in the fielded military capabilities of states.

[104] Asfandyar Mir, "What Explains Counterterrorism Effectiveness? Evidence from the US Drone War in Pakistan," *International Security* 43, No. 2 (2018): 45-83.; Asfandyar Mir and Dylan Moore, "Drones, Surveillance, and Violence: Theory and Evidence from a US Drone Program," *International Studies Quarterly* 63 No. 4 (2019): 846-862.; Aqil Shah, "Do U.S. Drone Strikes Cause Blowback? Evidence from Pakistan and Beyond," *International Security* 42, No. 4 (2018): 47-84.

Chapter 16: Uncrewed Aircraft Technology ca. 2010-2020

Thus far, this dissertation has covered one hundred years of advances in uncrewed aircraft technology without finding strong evidence in support of the proposition that technological change tends to engender sizeable and abrupt shifts in the fielded military capabilities of states. To conclude Part Three of this dissertation, I will examine advances in LRFW uncrewed aircraft technology in the ca. 2010-2020 timeframe and explain why these did not result in a situation in which uncrewed aircraft of this type could increasingly substitute for, let alone supplant, crewed military aircraft in a wider range of roles beyond ISR.

Unlike earlier timeframes in which the number of notable LRFW uncrewed aircraft designs that existed beyond the drawing board could often be counted in the single digits, the dynamics of diffusion and emulation, as well as efforts by a growing number of states to advance LRFW uncrewed aircraft technology, resulted in an overabundance of LRFW uncrewed aircraft designs in the ca. 2010-2020 timeframe. Given the objective of discerning whether there is empirical support for the proposition that advances in LRFW uncrewed aircraft technology engendered sizeable and abrupt shifts in the fielded military capabilities of states in this timeframe, I will first briefly summarize developments in three groupings of LRFW uncrewed aircraft to both provide evidence of the enduring emphasis on ISR roles and explain the reasons for this continued emphasis. I will then briefly examine the dynamics of countering (i.e., the adoption of countermeasures) that increasingly moderated the effects of advances in LRFW uncrewed aircraft technology on fielded military capabilities in this timeframe.

LRFW uncrewed aircraft technology diffused widely between ca. 2010 and 2020 following the high-profile and large-scale employment of such aircraft by the United States in both unarmed and armed ISR roles over the prior decade. The United States and Israel, the leaders in LRFW uncrewed aircraft technology at the turn of the decade, found a growing number of export customers, and this dynamic of diffusion was amplified by the development of competitive and broadly comparable uncrewed aircraft by a growing number of other countries that, in time, exported LRFW uncrewed aircraft of their own.[105] As a result, whatever the effects of LRFW uncrewed aircraft technology on the fielded military capabilities of states, the moderating effects of diffusion resulted in a situation in which capability gains were less concentrated than these had been in the ca. 1990-2010 timeframe.

16.1 New Developments: LRFW Uncrewed Aircraft Weighing Under 1,000 Kilograms

Given the cost and complexity of both developing and operating LRFW uncrewed aircraft, most countries focused on designs that (A) were optimized for glide performance, which made the resulting aircraft more suitable for ISR than other roles, and (B) weighed under 1,000 kilograms, which is to say smaller aircraft that lacked the capacity to undertake long endurance missions and/or carry a large payload or operate at high altitudes but were nevertheless less complex and costly to build and operate. In addition to American and Israeli designs, Chinese, Iranian, South African, Italian, South

[105] Dan Gettinger, "The Drone Databook," (Annandale-on-Hudson, NY: Center for the Study of the Drone, Bard College, 2019), https://dronecenter.bard.edu/projects/drone-proliferation/databook/

Korean, and Turkish designs, among others (this is not an exhaustive listing of countries), drove the diffusion of LRFW uncrewed aircraft to a greater number of states. All things considered, these existing and new LRFW uncrewed aircraft weighing under 1,000 kilograms offered broadly comparable performance as aircraft (when compared with LRFW uncrewed aircraft of similar weight, engine power output, and wingspan), and these incorporated the latest advances in sensor and datalink technology (there was, however, considerable variance across countries in terms of the performance of onboard sensors). As a result, LRFW uncrewed aircraft developed and built in the 2010s offered higher-resolution sensor imagery incomparable with the lower-resolution imagery common on smaller, low-end electro-optical sensors used throughout the 1990s and early 2000s.

While advances in sensor and radio technology made LRFW uncrewed aircraft more capable in ISR roles, such aircraft would not be very competitive with crewed aircraft without advances in automation technology to avoid the longstanding problem of uncrewed aircraft crashing following disruptions to the line-of-sight radio command link. Progressive advances in aircraft automation technology resulted in a situation in which LRFW uncrewed aircraft designed and built in the 2010s were increasingly automated and increasingly reliable aircraft capable of maintaining level flight and returning toward the preprogrammed landing area in the event of the disruption of the line-of-sight radio link. Stated differently, while most LRFW uncrewed ISR aircraft in this timeframe were of the remotely piloted variety, the maintenance of a human in the loop was increasingly pursued to have a human in a supervisory capacity to make decisions in the context of a

mission and/or respond to sensory data in real-time rather than actively operating the flight controls of an aircraft that would otherwise lose control and crash.

The cumulative effect of this diffusion of LRFW uncrewed aircraft weighing under 1,000 kilograms was that a growing number of states benefitted from gains in their battlefield ISR capabilities. It is, however, important to recognize two dynamics that moderated the shifts engendered in fielded military capabilities.

First, the number of LRFW uncrewed aircraft, or any other type of equipment, employed by a military, strongly shapes the overall magnitude of effects on fielded capabilities. In the ca. 2000 to 2010 timeframe, the United States built and deployed hundreds of LRFW uncrewed aircraft and these complemented a similarly large fleet of crewed fixed- and rotary-wing aircraft of all types that progressively received the latest in sensor technology. In contrast, as late as 2020, perhaps no more than twenty countries operated more than two dozen LRFW uncrewed aircraft. The reasons for this disparity are, of course, varied and multifaceted, but the simplest explanation is that even LRFW uncrewed ISR aircraft weighing under 1,000 kilograms remained expensive single-role (i.e., ISR) assets for most militaries. As a result, militaries with a lower capacity to undertake large-scale procurement efforts either adopted such LRFW uncrewed aircraft in small quantities – often in the single digits – and/or allocated their limited resources toward smaller (including human-held) reusable uncrewed aircraft, whether fixed- or rotary-wing, that weighed under 100 kilograms (i.e., a category of uncrewed aircraft technology that is beyond the scope of this dissertation).

A second dynamic that moderated the shifts engendered in fielded military capabilities relates to the plateauing of uncrewed aircraft technology identified toward the

end of the previous chapter. All things considered, the LRFW uncrewed aircraft fielded in the 2010s were broadly comparable to other designs of similar weight, engine power output, and wingspan. There were, in other words, no substantial developments in the areas of airframes and propulsion systems of LRFW uncrewed aircraft, and technological advances were, as before, concentrated in the areas of sensors and radio datalinks and, to a lesser degree, onboard aircraft automation. To appreciate this dynamic, we must of course consider technological advances in the subcategory of LRFW uncrewed aircraft weighing over 1,000 kilograms as the pursuit of increased performance beyond that provided by upgrading an existing airframe with the latest in onboard aircraft automation, sensors, and radio datalinks invariably required larger and heavier – more complex and more expensive – LRFW uncrewed aircraft designs.

16.2 New Developments: LRFW Uncrewed Aircraft Weighing Over 1,000 Kilograms

While less complex and more affordable LRFW uncrewed aircraft weighing under 1,000 kilograms became increasingly prevalent during the ca. 2010-2020 timeframe, the cost and complexity of designing and manufacturing and/or procuring LRFW uncrewed aircraft weighing more than 1,000 kilograms, to say nothing of the costs associated with operating said aircraft, precluded the widespread diffusion of more capable LRFW uncrewed aircraft that could better compete with crewed aircraft for roles.

In addition to the export the of American MQ-1 and MQ-9 and the Israeli Heron, all of which were 1990s designs progressively updated to reflect the latest advances in onboard aircraft automation, sensors, and radio datalinks, other countries developed

heavier and higher-performance propeller-driven endurance-optimizing LRFW uncrewed ISR aircraft of their own. Israel developed – and exported – the Hermes 900, a design in the same performance class as the American MQ-1. China developed and fielded a wide array of LRFW uncrewed aircraft including the GJ-1/Wing Loong-I and the GJ-2/Wing Loong II. Equipped with a piston engine and a turboprop respectively, these Chinese LRFW uncrewed ISR aircraft became close analogs to the American MQ-1 and MQ-9, respectively. In addition, (other) Chinese defense companies also developed a range of other higher-performance LRFW uncrewed ISR aircraft such as the CH-4 and CH-5, which similarly paralleled the MQ-1 and MQ-9, respectively, as well as the BZK-005. Beyond their employment by the Chinese military, the Wing Loong I, Wing Loong II, CH-4, and CH-5 were exported in large numbers over the 2010-2020 timeframe. For their part, Iran and Turkiye developed and fielded broadly comparable aircraft of their own in the form of the Iranian Shahed-129 and the Turkish Anka. Other countries, including South Korea, Taiwan, Russia, and India, also undertook developmental efforts but these did not result in deployed systems by 2020.

Viewed in isolation, this dynamic of adoption and diffusion not only resulted in a growing number of states experiencing gains in their battlefield ISR capabilities but also resulted in the fielding of increasing numbers of uncrewed ISR aircraft that could better compete with crewed ISR aircraft for funding and roles. Nevertheless, all the aforementioned aircraft designs are reflective of the plateauing of LRFW uncrewed aircraft technology discussed in the previous chapter. While these uncrewed aircraft incorporated progressively more advanced onboard automation, sensors, and radio datalinks, the aforementioned designs remained close analogs to the MQ-1 (i.e., glide-

optimizing LRFW uncrewed aircraft of a given weight class equipped with a piston engine) and the MQ-9 (i.e., glide-optimizing LRFW uncrewed aircraft of a given weight class equipped with a turboprop engine) in terms of aircraft performance. There was, in other words, very little in the way of advancement in the airframes of fielded LRFW uncrewed aircraft in the ca. 2010 to 2020 timeframe. To further optimize for long endurance operations, successively larger and heavier LRFW uncrewed ISR aircraft designs invariably received longer wingspans and some aircraft received more powerful turboprop engines – in some cases two engines – to increase the payload of the aircraft but there was little if any growth in terms of aircraft maneuverability and speed, among other aircraft performance parameters. Being restricted to high aspect ratio wings to obtain long endurance and being reliant on commercially available engines for efficient fuel consumption at the lower speeds necessary to sustain high endurance without carrying a much greater – heavier and larger – fuel load, the latest specimens of LRFW uncrewed aircraft technology could not match the capabilities offered by many types of crewed military aircraft.

The LRFW uncrewed aircraft weighing over 1,000 kilograms developed and deployed in the ca. 2010-2020 timeframe were, as a result of progressive advances in sensors and radio datalinks, much more capable in the undertaking of ISR roles than earlier designs, but the extent of performance growth in terms of aircraft design delineated the scope of roles for which these LRFW uncrewed aircraft were competitive. To be clear, the increased payload of the larger, heavier, and more complex aircraft offered an increased capacity to undertake armed ISR roles, but this was not without cost or trade-offs, and, no less importantly, was of greater relevance in more permissive

military-geographical contexts than those in which adversaries operated competitive crewed combat aircraft and air defenses. All things equal, the greater the weight of munitions carried under the wings, the greater the aerodynamic drag, the greater the fuel consumption, and the lower the amount of fuel that could be carried onboard, thereby resulting in reduced maximum range and endurance than that which could be obtained in an unladen "clean" flying configuration. In the context of armed ISR operations over permissive airspace characterized by an absence of air defenses, these larger, heavier, and more expensive LRFW uncrewed aircraft offered militaries considerable capability gains. To better compete with crewed aircraft for funding and roles in less permissive airspace, however, militaries required substantial advances in the areas of airframes, propulsion, and, not least, onboard aircraft automation to develop higher-end LRFW uncrewed aircraft that could realistically undertake missions in contested airspace and return to undertake additional missions and thereby earn the designation of large *reusable* fixed-wing uncrewed aircraft in such military-geographical contexts.

16.3 New Developments: Higher-End LRFW Uncrewed Aircraft

To better compete with crewed military aircraft for funding and roles, the designers of LRFW uncrewed aircraft had three (non-exclusive) options for higher-end designs. First, they could develop far larger, heavier, and more expensive LRFW ISR aircraft and essentially emulate the American jet-powered RQ-4. By doing so, designers could develop high altitude long endurance uncrewed ISR aircraft capable of carrying a payload that could match that of all but the most sophisticated crewed ISR aircraft such as the U-2. Second, they could pursue an intermediate aircraft configuration that offered greater

speed and payload through the use of jet propulsion but retained the emphasis on flight endurance and thereby remained primarily oriented toward ISR and armed ISR roles. Third, they could develop LRFW uncrewed aircraft that dispensed with high aspect ratio wings and the optimization on glide performance required to undertake long endurance missions and instead make use of progressive if gradual advances in aircraft automation to undertake missions other than ISR and thereby develop LRFW uncrewed aircraft that could more credibly compete with crewed combat aircraft for funding and roles.

16.3.1 Larger, Heavier, and More Expensive High Altitude Long Endurance LRFW Uncrewed Aircraft

The first option, which entailed developing aircraft along the lines of the American RQ-4, offered the possibility of carrying sensor payloads of greater weight and volume that could better compete with the sensor payloads found on higher-end crewed ISR aircraft albeit in an airframe that was (A) not subject to human physiological limitations and (B) did not put any human operator at risk. To attain such performance, aircraft like the RQ-4 had to be equipped with more powerful jet engines that were efficient when operating at higher altitudes and higher speeds than propeller-driven aircraft. As a result, these aircraft had to carry more fuel to match the endurance of slower propeller-driven aircraft and were considerably larger, heavier, and more expensive than even the likes of the turboprop-powered MQ-9 and its non-American counterparts. The RQ-4, for example, had a maximum take-off weight of around 15,000 kilograms, which is about 75% of the maximum take-off weight of a (fully loaded) crewed F-16 fighter, 80% of the maximum take-off weight of the crewed U-2 ISR aircraft, 300% of the maximum take-off weight of

the uncrewed turboprop-powered MQ-9 Reaper, and 1500% of the maximum take-off weight of the uncrewed piston-engine equipped MQ-1 Predator.

The cost and complexity of designing and operating such high-end LRFW uncrewed ISR aircraft resulted in just four countries well as NATO as a collective joining the United States in operating the RQ-4 and a maritime patrol variant of this design, the MQ-4C, in small, single-digit quantities. As a result of the failure of a multinational European effort to develop the smaller jet-powered EADS Talarion, only one country, China, followed the United States in developing an LRFW uncrewed ISR aircraft optimized for high altitude long endurance operations in the ca. 2010-2020 timeframe. While credible information on the Chinese WZ-7 Soaring Dragon is limited, satellite imagery showing multiple copies of the aircraft deployed at airbases near the South China Sea and in the Himalayas, imagery from aerial intercepts released by the Japanese military, and the appearance of the aircraft at the 2022 Zhuhai Airshow all indicate that it is a large, heavy jet-powered aircraft with an airframe optimized for high altitude long endurance missions along the lines of the RQ-4.

At first glance, the likes of the RQ-4 and WZ-7 indicate that leading militaries could, after making the necessary resource allocations, develop and deploy high(er)-end LRFW aircraft capable of undertaking high altitude and long endurance ISR missions in a manner that allowed these uncrewed aircraft to better compete with crewed ISR aircraft for funding and roles. The reality, one reflected in the fact that very small numbers of such aircraft exist despite decades of technological development dating back to the 1960s, is very different. As late as 2020, larger, heavier, and more expensive high altitude long endurance LRFW uncrewed aircraft were more desirable in theory than in practice

and were, as such, constituted unaffordable luxury capability set other than in small numbers even for countries with the vast resources of the United States and China.

While the increased weight of the likes of the RQ-4 and WZ-7 largely reflected upon the larger payload for sensors and fuel, it also reflected the reality that the cost of the airframe and its high-end sensors and radio datalinks required much greater levels of automation such that the aircraft could operate without human supervision in the event of a disruption to the line-of-sight radio datalink. It was, after all, one thing for a military to lose a smaller, more affordable LRFW uncrewed aircraft that cost several hundreds of thousands of dollars or even several million dollars as a result of the disruption to the line-of-sight radio datalink and a wholly different matter to lose a high-end LRFW uncrewed aircraft that cost tens of millions of dollars to manufacture (i.e., excluding large research and development costs that had to be amortized over a small number of high-end production aircraft). With the cost of manufacturing large, heavy LRFW uncrewed aircraft packed with high-end sensors approaching if not exceeding the cost of manufacturing large, heavy crewed aircraft (i.e., business jets and narrow-body passenger aircraft converted to undertake ISR roles), the likes of the RQ-4 and WZ-7 required much more advanced forms of onboard aircraft automation technology and this was anything but easy and inexpensive to develop, test, and implement on real, flying aircraft.

Although the advanced levels of automation found on the RQ-4 and, presumably, the WZ-7, resulted in high-end LRFW uncrewed ISR aircraft capable of *reliably* undertaking long-distance missions with minimal human intervention, the two inherent areas of comparative advantage of LRFW uncrewed aircraft over crewed aircraft – not being bound by human physiological constraints and not putting an onboard human

operator at risk – were of relatively limited consequence in practical terms during peacetime (i.e., in the absence of inter-state war) outside of counter-terrorism and counter-insurgency operations. To understand this dynamic, it is productive to compare the uncrewed RQ-4 with the crewed U-2.

The RQ-4, being uncrewed, could undertake long endurance missions that exceeded the physiological limitations of the crewed U-2's onboard pilot. The RQ-4, with a maximum take-off weight of approximately 80% that of the U-2, was, however, a considerably smaller aircraft, and the crewed U-2 could carry a heavier and larger load of sensors such that the uncrewed RQ-4 could not substitute for the U-2 in the entirety of the crewed aircraft's roles (the RQ-4 could, however, carry enough fuel to undertake very long endurance missions unlike the U-2). In peacetime, which is both the primary operational environment in which the undertaking of long endurance ISR missions by large and expensive aircraft is most relevant and a context in which human pilots operating near other countries' airspace are relatively safe, the *only* advantage of the RQ-4 over the U-2 was its capacity to remain airborne for twenty or more hours, something that human physiological limitations did not permit on the U-2. Of course, the United States could have developed a larger aircraft along the lines of the RQ-4 to match or possibly exceed the U-2 across all performance parameters including payload while remaining capable of undertaking long endurance missions but such a hypothetical aircraft would have to be substantially larger and heavier – and invariably more expensive to develop, build, and operate – than either the RQ-4 or the U-2. After all, the RQ-4's superior endurance was not without cost in terms of design trade-offs – the fuel load of the RQ-4, without which the RQ-4 could not exceed the U-2's endurance, came at

the cost of weight and volume and large amounts of fuel were required to remain airborne for twenty or more hours even in a high-altitude aircraft with excellent glide performance and an efficient turbofan engine.

In a world of infinite resources, militaries could simultaneously pursue the likes of the RQ-4 and a larger, more capable aircraft that could match or exceed the U-2's performance while undertaking long endurance missions. In the real world, militaries, including the amply funded American military, not only lacked the resources to develop and manufacture such aircraft in large numbers but also lacked the resources required to operate *orbits* of high-end ISR aircraft, crewed or uncrewed, on a sustained basis so as to take full advantage of the very long endurance of high-end LRFW uncrewed aircraft like the RQ-4 and WZ-7. In a world in which atmospheric ISR aircraft, crewed or uncrewed, compete with ISR satellites by offering the capacity to undertake short notice and more (relative to satellites in low earth orbit) persistent operations, militaries, including both the American and Chinese militaries, allocated their finite resources toward a broad range of ISR aircraft and capabilities, crewed and uncrewed (including orbital satellites), and the cost convergence of high-end LRFW uncrewed ISR aircraft such as the RQ-4 and WZ-7 with high-end crewed ISR aircraft meant that the latter, which offered greater mission flexibility and payload at the cost of putting human operators at risk – a relatively minor consideration in peacetime – and being incapable of operations much longer than ten hours, remained the preferred approach for most ISR roles while high-end uncrewed ISR aircraft such as the RQ-4 and WZ-7 were relegated to a niche subset of roles for which only small quantities of aircraft were required. As such, leading militaries worldwide operated far greater numbers of larger, higher-end crewed ISR aircraft than

larger, higher-end uncrewed ISR aircraft as late as 2020 and this dynamic was paralleled in terms of development and procurement programs for ISR aircraft, crewed and uncrewed, of all types.

16.3.2 More Capable Intermediate High Altitude Long Endurance LRFW Uncrewed Aircraft

High-end LRFW uncrewed ISR aircraft like the RQ-4 and WZ-7 were very expensive to develop, build, and operate and were, all things considered, generally not competitive with high-end crewed ISR aircraft in terms of cost or across a range of performance parameters. As a result, these were primarily restricted to niche roles for which only small numbers of aircraft were required, and the consequence of this dynamic was that (A) such uncrewed aircraft did not supplant crewed ISR aircraft and (B) the introduction of small numbers of niche high-end LRFW uncrewed ISR aircraft did not engender major gains in fielded military capabilities. Nevertheless, the limitations of the more prevalent propeller-driven LRFW uncrewed ISR aircraft in the vein of the MQ-1 and MQ-9 were well-known, and some (well-resourced) militaries remained interested in an intermediate category of LRFW uncrewed aircraft that used jet engines to obtain higher speed and greater payload without the cost and complexity of the likes of more exquisite high-end designs like the RQ-4 and WZ-7. If such aircraft could be developed, the thinking went, a new intermediate category of LRFW uncrewed ISR aircraft equipped with jet engines could substitute for smaller crewed ISR aircraft (i.e., ISR configured small business jets and medium-sized propeller-driven passenger aircraft) in a subset of ISR roles – and

potentially armed ISR roles – for which the likes of the MQ-1 and MQ-9 were not well suited.

As explained earlier, trends in technology development in the ca. 2010 to 2020 timeframe amounted to the plateauing of LRFW uncrewed aircraft technology in terms of airframes and this restricted the roles for which LRFW uncrewed aircraft could compete against crewed aircraft. Tellingly, during this timeframe, only two intermediate jet-powered LRFS uncrewed aircraft were developed, the American MQ-20 Avenger and the considerably smaller Chinese WZ-10. Based on publicly available information, both aircraft amount to jet-powered, higher payload, and moderately more automated analogs to the likes of the remotely piloted MQ-9 Reaper, which is to say that the onboard automation technology could operate without constant human supervision but could not undertake complex missions without the maintenance of a human operator in the loop through a radio datalink.

All things considered, the cost and complexity of larger, heavier, and more expensive higher-performance jet-powered LRFW uncrewed ISR aircraft was not offset by the limited gains offered in terms of capabilities, primarily in the areas of speed, altitude, and payload. Aircraft like the MQ-20 and WZ-7 compared unfavorably with heavier propeller-driven LRFW uncrewed ISR aircraft like in MQ-9 in terms of endurance, acquisition costs, and operating costs, and were more suited to niche roles for which fewer aircraft were required and competition with crewed ISR aircraft was greater. Stated differently, more capable intermediate high altitude long endurance LRFW uncrewed aircraft did not have much of a comparative advantage over crewed aircraft. For this reason, the American MQ-20 did not enter frontline service and publicly

available information indicates that the Chinese WZ-7 entered service on a limited scale alongside an array of LRFW uncrewed aircraft of various weight and performance classes (notably, China lacks the United States' large and diverse array of crewed ISR aircraft).

16.3.3 The Pursuit of LRFW Uncrewed Aircraft for Roles Beyond ISR

As explained in the preceding pages and the prior chapter, as late as 2020, essentially every LRFW uncrewed aircraft in service had a design optimized for long-endurance ISR roles. This was, as explained, no accident, for long endurance ISR roles were both an area in which LRFW uncrewed aircraft had an advantage over crewed aircraft and an area in which extant onboard automation systems were adequate. Absent further advances in onboard aircraft automation technology, LRFW uncrewed aircraft were not practical for complex missions not characterized by the relatively stable higher-altitude flight parameters used by LRFW uncrewed ISR aircraft.

These limitations in aircraft automation technology created a problem for militaries as LRFW uncrewed aircraft offered, in principle at least, the potential for major capability gains by not putting a human pilot at risk and avoiding human physiological limitations. While designers and militaries alike became increasingly skeptical of the near-term possibility of developing uncrewed aircraft that could fully substitute for crewed aircraft across a range of combat roles in the near term over the ca. 2010-2020 timeframe, developmental efforts refocused on new LRFW uncrewed aircraft designs that could both augment and complement – not supplant – crewed aircraft in combat roles.

By 2020, a growing number of militaries began to pursue two new types of LRFW uncrewed aircraft for roles beyond ISR and armed ISR. None of these two new types of aircraft entered service as of 2020 and the effects on the *fielded* military capabilities of the states undertaking such research and development efforts were, as such, zero. Nevertheless, a condensed summary of these technological developments allows us to understand what militaries did and did not hope to realize in the near to medium term and this better positions us to determine whether technological advances through 2020 were approaching a stage that allowed LRFW uncrewed aircraft to substitute for or supplant crewed aircraft in roles beyond ISR. If such evidence is to be found, then there may be a case to be made that although advances in LRFW uncrewed aircraft technology from ca. 1910 to 2020 did not engender sizeable and abrupt shifts in fielded military capabilities, major impending shifts may be underway. If, however, such evidence is not to be found, then the experiences of the ca. 2010 to 2020 timeframe provide further credence to the argument that the experience with advances in LRFW uncrewed aircraft technology has been no different than the experiences with submarine technology and, as such, do not lend support to the proposition that technological change tends to engender sizeable and abrupt shifts in the fielded military capabilities of states.

16.4 Renewed Effort Toward Developing LRFW Uncrewed Strike Aircraft

To pursue roles beyond ISR, militaries pursued two primary approaches to LRFW uncrewed aircraft technology in the 2010-2020 timeframe, both of which entailed dispensing with the longstanding airframe design optimization on long endurance

missions with high aspect ratio wings, and both of which were predicated upon substantial advances in aircraft automation technology.

The first new type of LRFW uncrewed aircraft (prototypes) pursued built upon efforts made earlier in the 2000s to develop uncrewed aircraft with low radar signatures which typically meant tailless aircraft (i.e., flying wing and blended wing aircraft). Such an uncrewed aircraft, the thinking went, could be used to undertake strike missions either independently or in conjunction with crewed combat aircraft. With guided bombs using a combination of inertial navigation and satellite-based radio navigation capable of functioning as "fire-and-forget" munitions from so-called standoff ranges, an LRFW uncrewed aircraft required a level of automation sufficient to reach a preprogrammed target area, launch the guided munitions toward the preprogrammed target(s), and return to base for rearming, refueling, and reuse. If such LRFW uncrewed aircraft could be developed, these could better compete with and possibly substitute for crewed combat aircraft in high-risk strike missions in contested airspace.

In the United States, the military developed and tested the X-47B, the first LRFW uncrewed strike aircraft to operate from an aircraft carrier, while a privately funded initiative by Boeing resulted in the Boeing Phantom Ray technology demonstrator. The United Kingdom, for its part, developed and tested the BAE Taranis technology demonstrator aircraft and multiple European countries joined a program led by France to develop and test the Dassault nEUROn technology demonstrator. China, in turn, developed and tested an aircraft later designated the Hongdu GJ-11, an aircraft that may have transitioned from a technology demonstrator to a production standard design intended for frontline service as of this writing. Russia, a laggard across LRFW uncrewed

aircraft technology, developed and tested the Sukhoi S-70, an aircraft that as of this writing appears to be part of an ambitious procurement program rather than a technology demonstration effort.

While several of the aforementioned designs were subscale technology demonstrators unsuitable for employment in combat operations, the most influential and highest-profile design, the American X-47B, is worth examining in greater detail because it highlights an important and often overlooked trend in LRFW uncrewed aircraft technology – convergence with crewed combat aircraft in terms of size, weight, and cost.

Developed on the basis of the X-47A, which was essentially a subscale technology demonstrator that first flew in 2003, the X-47B was intended to demonstrate the feasibility of an LRFW uncrewed aircraft optimized for strike – not long endurance ISR missions – missions that could operate from an aircraft carrier and thereby potentially substitute for crewed combat aircraft on high-risk strike missions and do so with an airframe that traded maneuverability and speed for range, thereby extending the reach of a carrier air wing (at least in terms of strike missions). Powered by a single powerful turbofan engine of the same type used on the (single-engine) crewed F-16 and the (twin-engine) F-15 fighters, the low radar signature X-47B was a long-range subsonic tailless blended-wing-body aircraft with a maximum takeoff weight of over 20,000 kilograms. The X-47B was not, in other words, a miniaturized aircraft as it approached crewed combat aircraft in terms of size, weight, and, not unimportantly, cost (for comparison, the X-47B had a maximum take-off weight that was approximately 2000% than that of the RQ-1/MQ-1 Predator and 420% than that of the MQ-9 Reaper). In return, the X-47B offered the American military the capacity to carry up to 2,000 kilograms of

guided munitions – all carried within a bomb bay of fixed volume – to a range of around 1,500 kilometers from an aircraft carrier and could thereby extend the reach of a carrier air wing in air-to-surface strike roles (the X-47B did not have an onboard air-to-air radar and, more generally, lacked any air-to-air combat capability).

In principle, strike-oriented uncrewed aircraft like the X-47B offered a path forward through which LRFW uncrewed aircraft technology could better compete with crewed combat aircraft for funding and roles and there was no reason that the carrier-borne X-47B or a lightened derivative of the design could not operate from airbases on land. In principle, such designs offered militaries what amounted to a reusable cruise missile of sorts – a platform for releasing equally accurate guided bombs without expending the expensive airframe, propulsion system, and onboard electronics (guidance, navigation, etc.) found on a long-range cruise missile – as well as a means of undertaking long-range strike missions in contested airspace without putting the lives of human pilots at risk. In practice, enduring limitations in aircraft automation technology limited the roles for which LRFW uncrewed aircraft of this type were both practical and competitive in prospective real-world conditions.

If strike-oriented LRFW uncrewed aircraft such as the X-47B were to be simply launched with commands to fly toward a predesignated target area characterized by a low threat from enemy combat aircraft and ground-based air defenses, release GPS/INS-guided bombs (which, unlike laser-guided bombs, are fire-and-forget systems that do not require a human in the loop), and thereafter return to base, then militaries could use less expensive LRFW uncrewed aircraft like the MQ-9 to the same effect. To operate in a military-geographical context characterized by such a potent threat from adversary

aircraft and/or adversary air defenses that the operation of crewed combat aircraft was a high-risk endeavor in the first place, which was by definition the only type of operational context in which the cost and complexity of a low radar signature (i.e., "stealth") aircraft, whether crewed or uncrewed, could be justified, aircraft automation technology had to be capable of fully substituting for all the tasks fulfilled by a trained human pilot including the capacity to adapt in response to rapidly changing circumstances, threats, and opportunities. As late as 2020, militaries were more than capable of designing an LRFW uncrewed aircraft that could, with minimal if any human control inputs, navigate toward a given area, undertake one or more tasks (i.e., collect intelligence and/or release munitions) over the said area, and return to base. The state of automation technology had not, however, reached a stage wherein the likes of the expensive X-47B, when encountering hostile enemy aircraft – the X-47B was not designed with any air-to-air capability – or air defenses, could adapt to rapidly changing circumstances beyond attempting to evade enemy fire. And if the X-47B could not successfully evade enemy fire and rapidly adapt its mission to strike its intended targets, this would undermine the purpose of allocating scarce resources for the development of such a complex and comparatively exquisite LRFW uncrewed aircraft design. Consequently, this nominally reusable 20,000-kilogram LRFW uncrewed aircraft was a poor substitute for a much less expensive expendable cruise missile (in the vein of the American Tomahawk) let alone a crewed combat aircraft with a bounded but nevertheless more capable onboard human "computer" better positioned to adapt to rapidly changing circumstances in real-time (also capable of employing so-called standoff guided munitions to reduce the threat exposure of the onboard pilot).

For these reasons, high-end LRFW uncrewed aircraft designed for strike operations against well-equipped adversaries – as opposed to ISR missions in highly permissive airspace – were more appealing in concept than in practice and militaries voted with their wallets when it came to allocating resources for both research and development and procurement. It bears emphasis that notwithstanding the many innovations and technological feats it encapsulated, the X-47B was ultimately a means of delivering 2,000 kilograms of guided munitions – contained within the finite volume of an internal bomb bay in order to maintain the low radar signature – to a range of around 1,500 kilometers, which is conveniently close to the reported maximum range of the American Tomahawk cruise missile (which is equipped with a 450-kilogram warhead). For comparison, a crewed F-16 fighter equipped with the same engine as the X-47B could carry over 7,000 kilograms of guided munitions – all under its wings and fuselage. And if the X-47B was intercepted by enemy crewed fighter aircraft, it would be defenseless when under attack by gunfire if not by missiles. The X-47B was, in other words, an extremely expensive means of delivering a relatively modest payload over long range. In principle, it could serve as a useful complement to low radar signature crewed combat aircraft such as the American crewed F-35 fighter operating in its vicinity, but the 20,000-kilogram uncrewed aircraft was hardly inexpensive, and other than range, it had no comparative advantage over the F-35 to which it would be tethered (given the F-35's air-to-air capability) in truly contested airspace for which a low radar signature airframe was even required. To be more than an expensive long-range "bomb truck" (and little else) and effectively substitute for crewed combat aircraft, the designers of the X-47B required major advances in onboard aircraft automation that had not come to pass during

the ca. 2010-2020 timeframe despite decades of sustained research and development activity.

The enduring limitations of aircraft automation led militaries to consider a more modest role for LRFW uncrewed aircraft technology toward the end of the 2010s, one that aimed to complement and augment, rather than substitute for or supplant, existing and prospective crewed combat aircraft. This second new type of LRFW uncrewed aircraft technology pursued by militaries was characterized, as of 2020, by (A) the absence of heavy emphasis on low radar signatures required to undertake independent strike missions in contested airspace and (B) an emphasis on developing relatively low(er) cost – so-called "attritable" – uncrewed aircraft to serve as an adjunct to – a means of augmenting – crewed combat aircraft. In this second formulation, which has variously been called manned-unmanned teaming and loyal wingman (these have more recently been rebranded as collaborative combat aircraft in the United States), an LRFW uncrewed aircraft operates in the vicinity of – within line-of-sight (radio) of – a crewed aircraft to serve in one or more roles including offboard sensor platform, offboard jammer, offboard munitions carrier, and so forth. Stated differently, over the 2010s the objective increasingly became the development of LRFW uncrewed aircraft to augment crewed aircraft in combat roles much as guided munitions augmented crewed aircraft for decades. While a growing number of countries exhibited interest in "loyal wingman" type LRFW uncrewed aircraft over the 2010s, as of 2020, only the United States and Australia were publicly known to have built working prototype designs while China, the United Kingdom, and a collaborative effort of France, Germany, and Spain were reported to have developmental programs underway.

While these new uncrewed aircraft designs offered the potential of engendering gains in fielded military capabilities, none were deployed by the end of 2020, and it remains to be seen whether this new type of LRFW uncrewed aircraft will work as advertised and have the envisaged effects on fielded military capabilities. After all, visions of an increasingly uncrewed air combat future can be dated to the 1990s if not earlier, and, after some thirty years of development and background technological advances, expectations have not matched reality. These "loyal wingman" type LRFW uncrewed aircraft designs are nevertheless relevant for the present purposes for three reasons. First, this is an area of technological advancement that is very much decades in the making. Second, militaries reoriented their focus to complementing and augmenting crewed combat aircraft – as opposed to substituting for and supplanting crewed combat aircraft in the manner of the X-47B – as a result of enduring limitations in aircraft automation technology. Third, despite decades of technological development, militaries did not find themselves in a position in the 2010s in which increased resource allocation would open a proverbial tap from which service-ready LRFW aircraft designs would readily follow. Instead, as of 2020, militaries found themselves in the familiar position of having to undertake years of expensive and uncertain research and development followed by extensive testing before they could begin to deploy new military technologies that had the potential of engendering shifts in fielded military capabilities. As a result, the latest ongoing areas of technological development do not offer any indication that advances in LRFW uncrewed aircraft technology support the notion that technological change is suitable for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states.

16.5 The Observed Vulnerabilities of LRFW Uncrewed Aircraft to Countermeasures

Prior to this section focusing on the ca. 2010-2020 timeframe, there was little reason to discuss the effectiveness of countermeasures to LRFW uncrewed aircraft because this was a largely theoretical and hypothetical matter. It was, therefore, best understood – inferred – by examining the limitations of LRFW uncrewed aircraft relative to crewed combat aircraft and, to a lesser degree for the purposes of this dissertation, by examining advances in air defense capabilities (a complex and expansive topic that is beyond the scope of this dissertation) and the implications of said advances on crewed aircraft which, being composed of matter, have the same intrinsic vulnerabilities to detection and damage as uncrewed aircraft, all things equal. While American and NATO LRFW uncrewed aircraft were shot down by Serbian forces (including at least one RQ-1 Predator shot down by ground-based air defenses), and by Iraqi forces before the 2003 American-led invasion (including one RQ-1 Predator shot down by an Iraqi crewed combat aircraft), LRFW uncrewed aircraft were primarily employed in permissive military-geographical contexts against adversaries with limited if not non-existent air defense capabilities during the ca. 1990-2010 timeframe.[106] During the ca. 2010-2020 timeframe, however, LRFW uncrewed aircraft were used in a wider range of less materially lopsided conflicts, and this resulted in the loss of large numbers of LRFW uncrewed aircraft to enemy action.[107]

---

[106] For a relatively early study on the vulnerabilities of LRFW uncrewed ISR aircraft when employed against adversaries with air combat and/or air defense capabilities, see Andre Haider, *Remotely Piloted Aircraft Systems in Contested Environments: A Vulnerability Analysis* (Kalkar, Germany: Joint Air Power Competence Centre, 2014)

[107] "Drone Crash Database" *Drone Wars*, No Date, https://dronewars.net/drone-crash-database/

The civil wars in Yemen (from 2014) and Libya (after 2011, i.e., the Second Libyan Civil War), both of which included foreign intervention and combat operations involving state militaries and the heavily armed proxies of intervening states, constituted some of the first conflicts in which large numbers of LRFW uncrewed aircraft were used in sustained campaigns against adversaries that were (A) equipped with air defense capabilities and/or (B) also equipped with LRFW uncrewed aircraft.

During the Syrian Civil War, which was perhaps the defining conflict of the decade, the employment of LRFW uncrewed aircraft was primarily limited to small numbers of Iranian and Russian aircraft (operating in support of Syrian government forces) and much larger numbers of American and British aircraft (primarily operating against the Islamic State group in both Syria and northern Iraq), all of which were primarily employed against adversaries with very limited air defense capabilities.[108] In contrast, during the Yemeni Civil War, large numbers of primarily American and Chinese-built LRFW uncrewed aircraft operated by the American, Saudi, and Emirati militaries were employed against the coalition of Yemeni rebel forces led by Ansarallah, an armed group that accessed both the remnants of the air defense arsenal used by the erstwhile Yemeni government and some air defense systems supplied by Iran. Over the course of the war, twenty or more LRFW uncrewed ISR aircraft, including larger, higher-flying designs (i.e., weighing more than ~1,000 kilograms) operating in both ISR and armed ISR roles, were shot down by Ansarallah's ground-based air defenses, which were primarily limited to shorter range and less sophisticated systems that relied upon the

[108] Dan Gettinger, "Drones Operating in Syria and Iraq" (Annandale-on-Hudson, N.Y.: Center for the Study of the Drone, Bard College, December 2016), https://dronecenter.bard.edu/files/2016/12/Drones-in-Iraq-and-Syria-CSD.pdf

visual detection of airborne targets rather than ground-based radar (the Saudi-led coalition aggressively largely neutralized Yemen's modest radar network at the outset of its intervention in the Yemeni Civil War).[109] The conflict also involved the large-scale employment of Iranian-built uncrewed aircraft by Ansarallah in a single-use strike role (i.e., in a manner akin to a cruise missile) against targets in Saudi Arabia. Saudi Arabia claimed successful intercepts of many such uncrewed aircraft by its air defenses and combat aircraft, thereby providing further evidence of the vulnerabilities of LRFW uncrewed aircraft technology when employed in contexts in which adversaries had even a modicum of air defense capability, a characterization that could not be made of the U.S.-led conflicts against lightly equipped insurgents in Afghanistan and Iraq.

The Second Libyan Civil War also offered evidence of the vulnerabilities of LRFW uncrewed aircraft when operating in a less permissive operating environment. Whereas the conflict in Yemen was characterized by broad asymmetries in the equipment available to the opposing belligerents (i.e., the Ansarallah-led Yemeni rebels were not remotely as well equipped as the Saudi-led coalition that supported Ansarallah's Yemeni opponents, who almost exclusively served in ground combat roles), the conflict in Libya marked the first time that two belligerent groupings would deploy LRFW uncrewed aircraft in both ISR and armed ISR roles against one another. By 2020, Turkish intervention in the conflict led to the employment of Turkish Bayraktar TB2 LRFW uncrewed aircraft in both ISR and armed ISR roles in support of the Tripoli-based

[109] Garrett Reim, "Record Number of UAV Shoot Downs Prompt New USAF Tactics and Countermeasure Pod," *Flight Global*, June 30, 2020, https://www.flightglobal.com/military-uavs/record-number-of-uav-shoot-downs-prompt-new-usaf-tactics-and-countermeasure-pod/138908.article ; Joseph Trevithick, "Houthi Rebels Shoot Down U.S. Air Force MQ-9 Reaper Over Yemen," The War Zone, June 29, 2019, https://www.twz.com/14806/houthi-rebels-shoot-down-u-s-air-force-mq-9-reaper-over-yemen

Government of National Accord.[110] At the same time, Emirati intervention in the conflict in support of the Tobruk-based Libyan National Army led to the employment of Chinese-built Wing Loong II LRFW uncrewed aircraft in both ISR and armed ISR roles. To counter the other side's LRFW uncrewed ISR aircraft, the competing foreign intervenors deployed mobile air defense systems and Turkiye also deployed electronic warfare systems that subjected LRFW uncrewed aircraft to countermeasures not available to the combatants against whom American ISR and armed ISR uncrewed aircraft were primarily employed over the ca. 2000-2020 timeframe. As a result, both sides incurred significant losses of the relatively small number of LRFW uncrewed aircraft deployed in the conflict.[111]

Before these conflicts in the 2010s, the vulnerabilities of LRFW uncrewed ISR aircraft were largely a theoretical concern for which there was little empirical evidence. By 2020, however, the increasingly widespread employment of LRFW uncrewed aircraft technology culminated in the use of LRFW uncrewed ISR aircraft against adversaries with meaningful if limited air defense capabilities, thereby demonstrating the unsuitability of such uncrewed aircraft for use against other states, the enduring limitations of uncrewed aircraft technology despite decades of technological development, and the fact that LRFW uncrewed aircraft were not ready to supplant

[110] Franz-Stefan Gady, "Useful, but Not Decisive: UAVs in Libya's Civil War," International Institute for Strategic Studies, November 22, 2019, https://www.iiss.org/blogs/analysis/2019/11/mide-uavs-in-libyas-civil-war; Aaron Stein, "Say Hello to Turkey's Little Friend: How Drones Help Level the Playing Field," War on the Rocks, June 11, 2021, https://warontherocks.com/2021/06/say-hello-to-turkeys-little-friend-how-drones-help-level-the-playing-field/

[111] Chris Cole and Jonathan Cole, "Libyan War Sees Record Number of Drones Brought Down to Earth," Drone Wars, July 1, 2020, https://dronewars.net/2020/07/01/libyan-war-sees-record-number-of-drones-brought-down-to-earth/

crewed military aircraft even in niche ISR role outside of permissive operating environments.

Chapter 17: Reflecting on the Implications of Advances in LRFW Uncrewed Aircraft Technology ca. 1910 to 2020

In Chapter Three, I argued that an array of monadic and dyadic factors and considerations indicated that technological change is an unsuitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. In Part Two of this dissertation, which examined the effects of advances in submarine technology on fielded military capabilities over the ca. 1800-1945 timeframe, I explained that although technological change had non-zero effects on fielded military capabilities, a holistic assessment of this 145 or so year long experience does not present strong evidence in support of the notion that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in fielded military capabilities. The findings of Part Three of this dissertation, which have examined the implications of advances in uncrewed aircraft technology on fielded military capabilities over the ca. 1910-2020 timeframe, parallel those of Part Two and thereby further substantiate the claims made in the analytical framework developed in Chapter Three.

In Chapter Twelve, I anchored the history of uncrewed aircraft technology, broadly defined, in the early 1910s, and explained that major technological limitations restricted uncrewed aircraft to two roles – that of a target aircraft and that of a single-use strike munition – for several decades, thereby constraining the effects that advances in uncrewed aircraft technology could have on fielded military capabilities. By the early 1940s, however, the cumulative effects of some three decades of technological development facilitated the development of new forms of uncrewed aircraft which, in their initial form, were essentially remotely piloted aerial bombs, and this expanded the

roles for which uncrewed aircraft could be conceivably employed. While a wide range of uncrewed aircraft, broadly defined, were developed and deployed during the Second World War primarily to augment crewed aircraft, the effects on the fielded military capabilities of states were decidedly marginal in the context of an exceptionally high-intensity protracted conflict but wartime technological developments nevertheless laid the foundations for more impactful developments in the post-war era.

As explained in Chapter Thirteen, the Cold War created conditions in which uncrewed aircraft technology, broadly defined, became a major area of research and development activity. This resulted in the development and later deployment of, among other things, guided ballistic and cruise missiles, highly automated crewed interceptor aircraft, surface-to-air missile systems, and anti-ship missiles. By the late 1950s, the unique requirements of the United States for new means of undertaking peacetime aerial reconnaissance missions in adversary airspace resulted in the development of the first uncrewed aircraft, as well as the first large reusable fixed-wing uncrewed aircraft, employed in the intelligence, surveillance, and reconnaissance role. As a result, LRFW uncrewed aircraft technology from the 1950s onward was increasingly pursued to substitute for, and possibly supplant, crewed aircraft in ISR roles, and uncrewed aircraft were, as such, increasingly competing with crewed aircraft, among other military systems, for funding and roles. Although these research and development efforts progressively expanded the technological frontier, capability gains lagged considerably behind technological change, and as late as 1970, after more than a half-century of technological development, uncrewed aircraft remained possible but not very practical and, no less importantly, were uncompetitive when compared with crewed aircraft in

most roles, including the very intelligence, surveillance, and reconnaissance role with which late twentieth century and early twenty-first-century uncrewed aircraft would become so closely associated.

While uncrewed aircraft technology, broadly defined, advanced considerably over the ca. 1945-1970 timeframe covered in Chapter Thirteen, it was not until the 1970s that successive and sustained technological advances gradually resulted in more practical and more competitive uncrewed aircraft designs, including those suitable for intelligence, surveillance, and reconnaissance roles. With military technology, and the effects of advances in military technology, not existing in a vacuum, however, the implications of progressive advances in uncrewed aircraft technology over the ca. 1970-1990 timeframe covered in Chapter Fourteen were strongly shaped by the operational requirements and military-geographical context for which uncrewed aircraft designs were intended. Consequently, the effects were greater on the fielded military capabilities of a small number of countries, namely Israel, and the implications of technological change on fielded military capabilities therefore remained limited up until the end of the Cold War despite nearly eighty years of technological development. Although uncrewed aircraft, broadly defined, of other types, most prominently guided munitions of various configurations, were more widely employed in the late Cold War period and thereby progressively if gradually reshaped military balances, developmental efforts were mostly oriented toward augmenting and complementing crewed aircraft rather than substituting for let alone supplanting crewed aircraft in all roles other than ISR over the ca. 1970-1990 timeframe. Given the backdrop of increasingly automated crewed military aircraft and a wide range of uncrewed aircraft, broadly defined, employed in the late Cold War

period, this dynamic reflected not major opposition from militaries to uncrewed aircraft that could more practically substitute for, and possibly supplant, the latest crewed aircraft designs, but the enduring reality that uncrewed aircraft technology had major limitations and was very far from a state of maturity.

As explained in Chapter Fifteen, although the end of the Cold War resulted in the decline of military expenditures in most countries, it was also followed by conflicts that greatly lowered the bar for a practical uncrewed aircraft capable of undertaking ISR missions as a substitute for a crewed aircraft while simultaneously driving surging demand for aerial ISR capabilities for which uncrewed aircraft were uniquely well suited. Uncrewed aircraft, after all, intrinsically have no human operators onboard to be put at risk during ISR operations, and uncrewed aircraft are not therefore bound by human physiological constraints. With military requirements unexpectedly lowered, the post-Cold War conditions allowed for the development of uncrewed ISR aircraft capable of undertaking long-endurance missions, which is to say the development of large reusable fixed-wing uncrewed aircraft that could outcompete crewed aircraft in certain roles and thereby substitute for if not supplant crewed aircraft in such roles. These non-technological dynamics, coming on the heels of around half a century of sustained research and development activity that began during the Second World War, resulted in a situation in which uncrewed aircraft technology had progressively greater implications on the fielded military capabilities of states albeit primarily in highly asymmetrical dyads in terms of air combat and air defense capabilities. This dynamic reflected the fact that enduring technological limitations largely restricted large reusable fixed-wing uncrewed aircraft technology to ISR and, in time, armed ISR roles, as well as the fact that even the

latest uncrewed ISR aircraft were highly vulnerable to air defenses. The effects of technological change on fielded military capabilities were, therefore, non-zero over the ca. 1990-2010 timeframe but were nevertheless highly contingent on post-Cold War dynamics unrelated to the trajectory of uncrewed aircraft technology.

While uncrewed ISR aircraft, and later uncrewed armed ISR aircraft, played a progressively more prominent and more impactful role in military operations against primarily non-state adversaries with negligible if not non-existent air defenses in the first decade of the twenty-first century, the effects of technological change on military balances featuring materially comparable states were far more muted. This reflected less a lack of interest on the part of militaries for uncrewed aircraft designs that could undertake a wider range of roles against state adversaries and thereby substitute and possibly supplant crewed aircraft in roles not limited to ISR, but the reality that the limitations of onboard automation technology severely curtailed the roles for which because uncrewed aircraft were practical and competitive relative to crewed aircraft substitutes. As a result, uncrewed aircraft technology experienced a state of plateauing by circa 2010, at around the same time that the operational employment of uncrewed aircraft in ISR and armed ISR roles in permissive airspace increasingly captured the attention of scholars, policymakers, and the wider public. The effects of technological change on the fielded military capabilities of states were therefore not only highly contingent on non-technological factors, but the effects of technological change on fielded military capabilities were also not subject to boundless growth.

As explained in Chapter Sixteen, militaries and designers were aware of this dynamic of plateauing and undertook extensive research and development activity with

the aim of expanding the technological frontier and thereby facilitating the emergence of large reusable fixed-wing uncrewed aircraft that could more practically compete with crewed aircraft for funding and roles beyond ISR and, ideally, substitute for and possibly supplant crewed aircraft in a variety of roles. Although such efforts resulted in considerable expansions of the technological frontier over the 2010s, the effects on the fielded military capacities of states were minimal as most of these efforts reached no further than the testing phase by 2020. As with earlier timeframes, the effects of technological change lagged well behind the expansion of the technological frontier. By 2020, which is to say after more than a century of technological development, a future in which uncrewed aircraft could practically substitute for crewed aircraft in a wide range of roles remained a prospective development far on the horizon.

All things considered, the experience with advances in uncrewed aircraft technology over the ca. 1910-2020 timeframe – an approximately 110-year-long experience if one starts with the first autopilot systems and an 80 or so year-long experience if one starts with the functional remotely piloted target aircraft – undoubtedly offers evidence of the non-zero effects of technological change on fielded military capabilities but does not constitute strong evidence in support of the notion that technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. Notably, the experience with uncrewed aircraft technology parallels the experience with submarine technology including both the dynamic of protracted evolutionary development in which a new technology is possible but not practical and therefore has negligible effects on the

fielded military capabilities of states, and the dynamic of plateauing in which the effects of technological change on the fielded military capabilities of states are non-linear.

Conclusion: Reflecting on The Relationship Between Technological Change and Shifts in the Fielded Military Capabilities of States

This dissertation project set out to investigate the relationship between technological change and shifts in fielded military capabilities with a specific focus on whether technological change is a suitable candidate for the role of a prime mover engendering sizeable and abrupt shifts – gains – in the fielded military capabilities of states. As explained in Chapter One and Chapter Two, influential segments of international relations scholarship require *something* to drive shifts in fielded military capabilities and are predicated upon the notion that shifts in fielded military capabilities are of consequence for international politics. With the fielded military capabilities of states not being mere abstractions, international relations scholars must identify and explicate the material basis of shifts in fielded military capabilities of states or else perceptions of military capabilities, and as such perceptions of military threats whether contemporary or historical, would be indistinct from a hallucination. Many of the candidates for the role of a prime mover identified in international relations scholarship, however, are unsuitable for engendering shifts in fielded military capabilities that are both sizeable and abrupt.

The real-world military effects of monadic, that is nominal, shifts in fielded military capabilities, whether caused by technological change or any other factor, are necessarily co-produced with the actions taken and forgone by other states that have incentives and opportunities to blunt if not negate the military effects of said shifts and thereby avert a significant decline in their fielded military capabilities. While the ultimate magnitude, whether in monadic or dyadic terms, of shifts engendered in fielded capabilities is important, states and policymakers must be considerate of the timeframes

over which shifts in fielded military capabilities unfold. All things equal, if shifts in fielded military capabilities engendered by technological change accrue gradually, then states have more time to respond, and the technological advance in question may have more limited effects on the military balance and, as such, less acute military significance for international politics. If, however, the gains in fielded military capabilities engendered by technological change accrue rapidly to the state introducing the new technology, then technological change may have abrupt effects on the military balance and more acute military significance for international politics.

In principle, technological change, broadly defined, is an appealing candidate for the role of a prime mover engendering sizeable and abrupt shifts in the fielded military capabilities of states. There are direct mechanisms through which technological change, including advances in military technology, can translate into shifts in fielded military capabilities, technological change can also unfold quickly and, not least, technological change is regularly invoked in influential segments of international relations scholarship, among other areas, as one of the core contributors to shifts in the fielded military capabilities of states. As explained in the analytical framework developed in Chapter Three, however, multiple monadic and dyadic factors undermine the suitability of technological change for the role of a prime mover engendering shifts in fielded military capabilities that are both sizeable and abrupt. By disaggregating the effects of technological change on fielded military capabilities into several phases – development, testing, production, and deployment – Chapter Three highlighted how tenuous and contingent the military effects of technological change can be. Even in a purely monadic framework in which no adversaries exist to attempt to blunt if not negate the effects of

technological change on a military balance, these phases constitute bottlenecks regulating the throughout – both the pace and magnitude – of the effects of technological change on a military balance. In practical terms, it tends to take multiple years for a state to develop, test, manufacture, and deploy a new technology in such numbers as to have a significant effect on a military balance in a state dyad. As such, the mere existence of a new technology in production-ready form, let alone in prototype form, tends not to engender sizeable shifts in fielded military capabilities even though it may serve as a harbinger of such an eventuality.

The dyadic mechanisms examined in Chapter Three also highlighted how technological change has no intrinsic effects on a military balance as the effects are coproduced by the actions taken and forgone by an adversary even in a stylized two-actor model. All things equal, the more time that an adversary state has to respond, the greater the opportunities to adopt, counter, and offset the new technology and thereby blunt if not negate the effects of technological change on fielded military capabilities. In practice, other states tend to have considerable opportunities to undertake efforts, including the development and deployment of countermeasures, which can blunt if not negate the effects that technological change would otherwise have on their military capabilities in relation to those of their adversary.

For these reasons, international relations scholarship does not only require *something* to serve as a prime mover driving shifts in fielded military capabilities but also a candidate prime mover that can engender capability gains that are both sizeable and abrupt. While the analytical framework developed in Chapter Three suggests that technological change is not a suitable candidate for a prime mover that international

relations scholarship requires, Chapter Three was a purely analytical exercise and the findings it developed require empirical support. Part Two and Part Three of this dissertation closely examined the historical record with longitudinal assessments of the implications on fielded military capabilities of advances in submarine technology over the ca. 1800-1945 timeframe and advances in uncrewed aircraft technology over the ca. 1910-2020 timeframe. By tracing advances in these respective areas of technology alongside developments in countermeasures and thereby presenting evidence that advances in military technology tend not to engender sizeable and abrupt shifts in the fielded military capabilities of states, this dissertation offered empirical support for the analytical framework developed in Chapter Three.

With respect to submarine technology, Part Two of this dissertation presented evidence that submarine technology was subject to nearly a century of protracted evolutionary development during which military submarines were possible but not practical and, importantly, were uncompetitive when compared with other military systems. While submarine technology underwent a sustained growth spurt from the late 1890s onward, technological development began to plateau by the early 1910s as a result of long-running technological limitations. As a result, when states employed submarines of various design vintages in the First World War and subjected this ostensibly new military technology to real-world combat conditions, not only were submarines not as effective in sinking surface ships – the sole purpose for which states developed and fielded submarines – as many anticipated prior to the onset of war, but submarines proved susceptible to basic countermeasures that greatly blunted the effects of advances in submarine technology on fielded military capabilities. While an argument could be

made that submarines were prematurely employed in combat during the First World War, an argument that would be supportive of the central claims of this dissertation project, the extension of the longitudinal analysis into the Interwar Period in Part Two highlighted the enduring limitations of submarine technology, thereby attesting to the regularity of protracted technological development and the lagged effects of technological change on fielded military capabilities. Tellingly, the Second World War experience, which unfolded some 140 years after the first submarine underwent testing and over 40 years after the first practical self-propelled submarine underwent testing, further attests to the dynamic of continuity, with the very concept of a submarine as a submersible torpedo boat rendered increasingly untenable as the war progressed as a result of an intense measure-countermeasure competition. In 1945, as with prior decades, advances in submarine technology, however consequential over the long run, tended not to engender sizeable and abrupt shifts in the fielded military capabilities of states.

With respect to uncrewed aircraft technology, Part Three of this dissertation presented evidence that uncrewed aircraft technology was subject to over half a century of protracted evolutionary development during which uncrewed aircraft were possible but not practical and, importantly, were uncompetitive when compared with other military systems in most roles, including that intelligence, surveillance, and reconnaissance (ISR) role with which late twentieth century and early twenty-first-century uncrewed aircraft would be so intimately associated. While uncrewed aircraft technology, broadly defined, advanced considerably over the ca. 1910-1970 timeframe, it was not until the 1970s that sustained technological advances resulted in more practical uncrewed aircraft suitable for the ISR role. With military technology, and the effects of advances in military

technology, not existing in a vacuum, the implications of progressive advances in uncrewed aircraft technology over the ca. 1970-1990 timeframe were strongly shaped by the operational requirements and military-geographical context for which uncrewed aircraft designs were intended. As a result, the implications of technological change on fielded military capabilities remained limited up until the end of the Cold War.

The end of the Cold War was followed by conflicts that greatly lowered the bar for a practical uncrewed aircraft that could undertake ISR roles while simultaneously driving surging demand for aerial ISR capabilities. These dynamics, coming on the heels of around half a century of sustained research and development activity, resulted in a situation in which uncrewed aircraft technology had progressively greater implications on the fielded military capabilities of states albeit primarily in highly asymmetrical dyads in terms of air combat and air defense capabilities. This dynamic continued in the first decade of the twenty-first century and uncrewed ISR aircraft, including armed versions of such aircraft, played a prominent role in military operations against primarily non-state adversaries with negligible if not non-existent air defenses. At the same time, advances in uncrewed aircraft technology made in the ca. 2000-2020 timeframe had considerably more muted effects on military balances between materially comparable states less as a result of a lack of interest on the part of militaries for uncrewed aircraft designs that could undertake a wider range of roles against state adversaries and thereby substitute and possibly supplant crewed aircraft, but because the limitations of onboard automation technology severely curtailed the roles for which because uncrewed aircraft were practical and competitive relative to crewed aircraft substitutes. Hence, as late as 2020, which is to say after a century or so of technological development, a future in which

uncrewed aircraft could practically substitute for crewed aircraft in a wide range of roles remained a persistently prospective development far on the horizon. For these reasons, advances in uncrewed aircraft technology tended not to engender sizeable and abrupt shifts in the fielded military capabilities of states.

Given the preliminary findings derived from the empirical and analytical work undertaken in this dissertation project, international relations scholarship remains in search of a prime mover that lends itself to engendering sizeable and abrupt shifts in the fielded military capabilities of states. As explained earlier, without the identification and explication of the material basis of shifts in fielded military capabilities, perceptions of shifts in fielded military capabilities, and as such perceptions of military threats whether contemporary or historical, would be indistinct from a hallucination. With the findings being preliminary, further research will be required to test the validity of the analytical framework and to test the generalizability of the empirical work undertaken in this dissertation project.

Future research may, in part, examine the military effects of technological change in other ways and through different research methodologies. For example, future research may pursue a more, or less, restrictive conception of military effects than this dissertation project, which exclusively focused on shifts in fielded military capabilities. Research more oriented towards political and/or diplomatic effects may give greater weight to the development and testing of a new military technology than the manufacturing and fielding of a new military technology, and this may capture how technological development as a process affects international politics in potentially different ways than fully developed and fielded military equipment (i.e., the primary effect of technological

development may be the fear and apprehension caused rather than actual shifts engendered in fielded military capabilities). This could offer opportunities for archival research and the surveying of policymakers.

Future research may also pursue a different approach to examine the limited number of real-world cases in which most military technologies are tested in combat conditions representative of what these were designed for. This dissertation project purposefully avoided detailed examinations of operational dynamics, such as the experiences with submarines in the First World War and Second World War, out of concern about the degree to which generalizable inferences can be made from what were, all things considered, highly contingent and non-extrapolatable experiences. The approach undertaken in this dissertation is not, of course, without cost or trade-offs, but it nevertheless reflects an effort to control for the litany of confounding variables and thereby facilitate a longitudinal approach to assessing the effects of advances in military technology on fielded military capabilities across time and space. Future scholarship may identify more practical ways of integrating a longitudinal analytical approach with the highly fragmented records of real-world combat experiences.

While different directions are possible, future research can also continue the longitudinal assessments undertaken in Part Two and Part Three of this dissertation and examine other areas of military technology to discern the generalizability of the findings across time, space, and technology area. Although there is a general paucity of real-world cases and, even in peacetime, of major and protracted cases of high-intensity military-technological competition, there exist opportunities for cross-national studies that examine variance in the effects of technology change. In the area of submarine

technology, for example, there is considerable well-attested variation in Cold War experiences. Although the U.S.-Soviet dynamic, dominated as it was by nuclear propulsion and nuclear-armed submarines, has received substantial attention since the end of the Cold War (albeit almost entirely from an American perspective),[1] much of the Cold War experience in undersea warfare and naval technology more generally has been largely unexplored by international relations scholars and scholars in related fields, including military history. Developments in diesel-electric submarine technology since 1945, for example, have been almost entirely overshadowed by the experience with marine nuclear propulsion but outside the U.S.-Soviet dyad, advances in diesel-electric submarine technology, which were a continuation of the dynamics that were examined at the end of Part Two of this dissertation, had far greater implications for the military capabilities of most state dyads. Similarly, if scholars were to expand the scope of uncrewed aircraft technology to include developments in guided munitions much as this dissertation project did in Part Three, there is ample opportunity for examination of how different countries pursued automation, broadly defined, in aerospace in the post-1945 period and the implications of these varied approaches to technological development on their fielded military capabilities.

While there is considerable scope for different directions in future research on the military effects of technological change on international politics, it bears emphasis that although this dissertation's empirical work is nominally limited to detailed longitudinal case studies focusing on submarine and uncrewed aircraft technology, both of these areas of technological development lie at the intersection of a wide range of military and

[1] For a comprehensive account of Cold War dynamics, see Owen R. Cote, "The Third Battle: Innovation in the U.S. Navy's Silent Cold War Struggle with Soviet Submarines," *Newport Paper 16,* (Newport, RI: Naval War College, 2003)

civilian technologies. The case study on submarine technology over the ca. 1800-1945 timeframe found in Part Two, for example, did not merely focus on submarines and torpedoes but encompassed a wide range of developments in propulsion technology, radio technology, aircraft technology, and so forth. Similarly, the case study on uncrewed aircraft technology over the ca. 1910-2020 timeframe found in Part Three did not merely focus on uncrewed ISR aircraft – it also covered developments in guidance systems that facilitated advances in uncrewed aircraft technology, broadly defined, it dealt with a wide range of advances in communications technology, and so forth. This dissertation project's empirical work has, therefore, covered a wide range of discrete areas of military technology across time and space.

At a time of increasing international tensions, intensifying military-technological competition, and growing scholarly interest in historical, contemporary, and prospective advances in military technology, it is productive for international relations as a field to revisit its understanding of technological change in the military sphere. International relations scholars rely heavily on three literatures – the military effectiveness, military innovations literature, and offense-defense literature – for their understanding of the implications of advances in military technology, and these influential literature seldom delve into details of the technology development process, the testing process, the manufacturing process, the fielding process, the employment of fielded systems in variant operational conditions, or the coevolution of new military technologies alongside countermeasures. While there is growing scholarly interest in so-called “emerging

military technologies,"[2] and the implications, broadly defined, of technological change for international politics, scholarship tends to focus more on the political and diplomatic effects of technological change than the military effects. No less importantly, scholarship tends to overlook the fact that the military effects are coproduced in dyads featuring responsive and adaptive adversaries working to blunt if not negate the military effects of one another's efforts. Ultimately, while politics has primacy and although threat perceptions need not necessarily have a material basis rooted in fielded military capabilities, international relations scholars interested in shifts in the fielded military capabilities of states would benefit from further research in search of a prime mover and this dissertation project's preliminary findings offer a foundation that future research may productively build upon.

[2] See, for example, Caitlin Talmadge, "Emerging Technology and Intra-War Escalation Risks: Evidence from the Cold War, Implications for Today," *Journal of Strategic Studies* 42, No. 6 (2019): 864-887.; Jacquelyn Schneider, "The Capability/Vulnerability Paradox and Military Revolutions: Implications for Computing, Cyber, and the Onset of War," *Journal of Strategic Studies* 42, No. 6 (2019): 841-863.; Michael C. Horowitz, "Do Emerging Military Technologies Matter for International Politics?" *Annual Review of Political Science* 23 (2020): 385–400; Todd S. Sechser, Neil Narang, and Caitlin Talmadge, Editors, *Emerging Technologies and International Stability* (New York: Routledge, 2022)

Bibliography

Adams, Karen Ruth. "Attack and Conquer? International Anarchy and the Offense-Defense-Deterrence Balance." *International Security* 28, No. 3 (2003/2004): 45-83.

Adamsky, Dima P. "Through the Looking Glass: The Soviet Military-Technical Revolution and the American Revolution in Military Affairs." *Journal of Strategic Studies* 31, No. 2 (2008): 257–94.

Adamsky, Dima. *The Culture of Military Innovation: The Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the US, and Israel.* Stanford, CA: Stanford University Press, 2010.

Altmann, Jurgen and Frank Sauer. "Autonomous Weapon Systems and Strategic Stability." *Survival* 59, No. 5 (2017): 117–142.

Anthony, Alastair, Norman K. Walker, and Elizabeth DeSocio. *The Accuracy of the Azon Guided Bomb as Affected by Battle Conditions in World War II*. Bethesda, MD: Norman K. Walker Associates.

Armacost, Michael H. *The Politics of Weapons Innovation: The Thor-Jupiter Controversy*. New York: Columbia University Press, 1969.

Army Technology. "Phoenix Battlefield Surveillance UAV." Army Technology. June 8, 2000. https://www.army-technology.com/projects/phoenixuav/?cf-view.

Avant, Deborah D. *Political Institutions and Military Change: Lessons from Peripheral Wars*. Ithaca, NY: Cornell University Press, 1994.

Aviation Week. "U.S. Air Force Presses for Collaborative Combat Aircraft Cost Sharing." February 20, 2024. https://aviationweek.com/defense-space/aircraft-propulsion/us-air-force-presses-collaborative-combat-aircraft-cost-sharing/.

Barno, David and Nora Bensahel, *Adaptation Under Fire: How Militaries Changes in Wartime.* New York: Oxford University Press, 2020.

Bassett, Ross Knox. *To the Digital Age: Research Labs, Start-up Companies, and the Rise of MOS Technology*. Baltimore, MD: Johns Hopkins University Press, 2007.

Becker, Elizabeth, "Crisis in the Balkans: The Drones; They're Unmanned, They Fly Low, and They Get the Picture." *New York Times*, June 3, 1999. https://www.nytimes.com/1999/06/03/world/crisis-balkans-drones-they-re-unmanned-they-fly-low-they-get-picture.html.

Beckley, Michael. "Economic Development and Military Effectiveness." *Journal of Strategic Studies* 33, No. 1 (2010): 43-79.

Beckley, Michael. "The Power of Nations: Measuring What Matters." *International Security* 43, NO. 2 (2018): 7-44.

Bergen, Peter L. and Daniel Rothenberg, editors. *Drone Wars: Transforming Conflict, Law, and Policy*. Cambridge, UK: Cambridge University Press, 2014.

Bergin, Bob. "The Growth of China's Air Defenses: Responding to Covert Overflights, 1949-1974." *Studies in Intelligence* 57, No. 2 (2013): 19-28.

Biddle, Stephen and Ivan Oelrich. "Future Warfare in the Western Pacific: Chinese Anti-Access/Area Denial, U.S. AirSea Battle, and Command of the Commons in East Asia." *International Security* 41, No. 1 (2016): 7-48.

Biddle, Stephen, and Stephen Long. "Democracy and Military Effectiveness: A Closer Look." *Journal of Conflict Resolution* 48, No. 4 (2004): 525–46.

Biddle, Stephen, and Robert Zirkle. 1996. "Technology, Civil-Military Relations, and Warfare in the Developing World." *Journal of Strategic Studies* 19, No. 2 (1996): 171– 212.

Biddle, Stephen, and Stephen Long. "Democracy and Military Effectiveness: A Deeper Look." *Journal of Conflict Resolution* 48, No. 4 (2004): 525–46.

Biddle, Stephen. "Rebuilding the Foundations of Offense-Defense Theory." *Journal of Politics* 63, No. 3 (2001): 741-774.

Biddle, Stephen. "The Gulf War Debate Redux: Why Skill and Technology are the Right Answer." *International Security* 22, No. 2 (1997): 163-174.

Biddle, Stephen. "Victory Misunderstood: What the Gulf War Tells Us about the Future of Conflict." *International Securit*y 21, No. 2 (1996):139-179.

Biddle, Tami Davis. Rhetoric and Reality in Air Warfare: The Evolution of British and American Ideas About Strategic Bombing, 1914-1945. Princeton, NJ: Princeton University Press, 2002.

Black, Jeremy. *War and Technology*. Bloomington, IL: Indiana University Press, 2013.

Blackwelder, Donald I. "The Long Road to Desert Storm and Beyond: The Development of Precision Guided Bombs." Master's Thesis, Advanced Airpower Studies, Air University, 1992.

Blagden, David. "When Does Competition Become Conflict? Technology, Geography, and the Offense-Defense Balance." *Journal of Global Security Studies* 6, No. 4 (2021): 1-23.

Boot, Max. *War Made New: Technology, Warfare, and the Course of History, 1500 to Today*. New York: Penguin, 2006.

Borgiasz, William S. *The Strategic Air Command: Evolution and Consolidation of Nuclear Forces, 1945-1955*. Westport, CT: Praeger, 1996.

Boyle, Michael J. *The Drone Age: How Drone Technology Will Change War and Peace*. Oxford, UK: Oxford University Press, 2020.

Boyne, Walter J. "The Early Overflights." Air & Space Forces Magazine. June 1, 2001. https://www.airandspaceforces.com/article/0601overfly/.

Bracken, Paul. "Non-Standard Models of the Diffusion of Military Technologies." *Defense Analysis* 14, No. 2 (1998): 101–114.

Breemer, Jan S. "Defeating the U-boat: Inventing Antisubmarine Warfare." *Newport Paper 36*. Newport, RI: Naval War College, 2010.

Brodie, Bernard, and Fawn M. Brodie. *From Crossbow to H-Bomb: The Evolution of the Weapons and Tactics of Warfare*. Bloomington, IN: Indiana University Press, 1973.

Brodie, Bernard. *Sea Power in the Machine Age*. Princeton, NJ: Princeton University Press, 1941.

Brooks, Risa A. and Elizabeth A. Stanley, Editors. *Creating Military Power: The Sources of Military Effectiveness*. Stanford, CA: Stanford University Press, 2007.

Brooks, Risa A. *Shaping Strategy: The Civil-Military Politics of Strategic Assessment*. Princeton, NJ: Princeton University Press, 2008.

Brooks, Risa. "An Autocracy at War: Explaining Egypt's (in)Effectiveness in the 1967 and 1973 Arab-Israeli Wars." *Security Studies* 15, No. 3 (2006): 396-430.

Brooks, Risa. "An Autocracy at War: Explaining Egypt's Military Effectiveness, 1967 and 1973," *Security Studies* 15, No. 3 (2006): 396–430.

Brooks, Stephen G. and William C. Wohlforth. "The Rise and Fall of the Great Powers in the Twenty-First Century: China's Rise and the Fate of America's Global Position." *International Security* 40, No. 3 (Winter 2015-2016): 7-53.

Brose, Christian. "The New Revolution in Military Affairs: War's Sci-Fi Future." *Foreign Affairs* 98, No. 3 (2019): 122-134.

Brose, Christian. *The Kill Chain: Defending America in the Future of High-Tech Warfare*. New York: Hachette, 2020.

Brown, Louis. A Radar History of World War II: Technical and Military Imperatives. Bristol, UK: Institute of Physics Publishing, 1999.

Brugioni, Dino A. *Eyes in the Sky: Eisenhower, the CIA and Cold War Aerial Espionage*. Annapolis, MD: Naval Institute Press, 2010.

Bruning, John R. *Crimson Sky: The Air Battle for Korea*. London, UK: Brassey's, 1999.

Butterfield, Herbert. *History and Human Relations*. London, UK: Collins, 1951.

Buzan, Barry and Eric Herring, *The Arms Dynamic in World Politics*. Boulder, CO: Lynne Rienner Publishers, 1998.

Byman, Daniel L., and Matthew C. Waxman. "Kosovo and the Great Air Power Debate," *International Security* 24, Bo. 4 (2000): 5–38.

Calcara, Antonio, Andrea Gilli, Mauro Gilli, and Ivan Zaccagini. "Why Drones Have Not Revolutionized War: The Enduring Hider-Finder Competition in Air Warfare." *International Security* 46, No. 4 (2022): 130-171.

Calcara, Antonio, Andrea Gilli, Mauro Gilli, Raffaele, and Ivan Zaccagini. “Will the Drone Always Get Through? Offensive Myths and Defensive Realities.” *Security Studies* 31, No. 5 (2022): 791-825.

Cappella Zielinski, Rosella, and Ryan Grauer, “Organizing for Performance: Coalition Effectiveness on the Battlefield.” *European Journal of International Relations* 26, No. 4 (2020): 953-978.

Castillo, Jasen. *Endurance and War: The National Sources of Military Cohesion*. Stanford, CA: Stanford University Press, 2014.

Ceruzzi, Paul E. *GPS*. Cambridge, MA: The MIT Press, 2018.

Chicago Tribune. “Air Decoys Fooled Iraq, Navy Reveals.” September 19, 1991. https://www.chicagotribune.com/1991/09/19/air-decoys-fooled-iraq-navy-reveals/.

Clark, Bryan. *The Emerging Era in Anti-Submarine* Warfare. Washington, DC: Center for Strategic and Budgetary Assessments, 2015.

Cohen, David E., “Technical Report—The Mk-XIV Torpedo: Lessons for Today,” *Naval History Magazine* 6, No. 4 (1992).

Cohen, Eliot. “A Revolution in Warfare,” *Foreign Affairs* 75, No. 2 (March/April 1996): 37-54.

Coker, Christopher. *Warrior Geeks: How 21st Century Technology is Changing the Way We Fight and Think about War.* London, UK: Hurst, 2013.

Cole, Chris, and Jonathan Cole. “Libyan War Sees Record Number of Drones Brought Down to Earth.” Drone Wars. July 1, 2020. https://dronewars.net/2020/07/01/libyan-war-sees-record-number-of-drones-brought-down-to-earth/.

Collins, V. Lansing, “The Sultan’s Submarines.” Proceedings 100, No. 10. (1974).

Correll, John T. “The Emergence of Smart Bomb.” Air Force Magazine. March 2010. https://www.airandspaceforces.com/PDF/MagazineArchive/Documents/2010/March%202010/0310bombs.pdf.

Cote, Owen R. “The Third Battle: Innovation in the U.S. Navy’s Silent Cold War Struggle with Soviet Submarines.” *Newport Paper 16*. Newport, RI: Naval War College, 2003.

Cote, Owen. “The Politics of Innovative Military Doctrine: The U.S. Navy and Fleet Ballistic Missiles.” Ph.D. dissertation, Political Science, Massachusetts Institute of Technology, 1996.

Cox, Samuel J. “H-Gram 021: Operation Avalanche, Fritz X, and the Battle of Durazzo.” Naval History and Heritage Command, September 18, 2018. https://www.history.navy.mil/content/dam/nhhc/about-us/leadership/hgram_pdfs/H-Gram_021.pdf.

Curatola, John M. *Bigger Bombs for a Brighter Tomorrow: The Strategic Air Command and American War Plans at the Dawn of the Atomic Age, 1945-1950.* Jefferson, NC: McFarland & Company, 2016.

Dabrowski, Krzysztof. *Hunt for the U-2: Interceptions of Lockheed U-2 Reconnaissance Aircraft over the USSR, Cuba and the People's Republic of China, 1959-1968.* Warwick, UK: Helion and Company, 2020.

Day, Dwayne A. "Intersections in Real Time: the Decision to Build the KH-11 KENNEN Reconnaissance Satellite (part 1)." The Space Review, September 9, 2019. https://www.thespacereview.com/article/3791/1.

Day, Wayne A., John M. Logsdon, and Brian Latell, editors. *Eye in the Sky: The Story of the Corona Spy Satellites.* (Washington, D.C.: Smithsonian Institution Press, 1998).

de Havilland Aircraft Museum. "De Havilland DH82B Queen Bee." No Date. https://www.dehavillandmuseum.co.uk/aircraft/de-havilland-dh82b-queen-bee/.

deLeon, Peter. *The Laser-Guided Bomb: Case History of a* Development. Santa Monica, CA: RAND Corporation, 1974.

Demchak, Chris C. *Military Organizations, Complex Machines*. Ithaca, N.Y.: Cornell University Press, 1991.

Desch, Michael C. "Democracy and Victory: Why Regime Type Hardly Matters." *International Security* 27, No. 2 (2002): 5-47.

Desch, Michael C. *Power and Military Effectiveness: The Fallacy of Democratic Triumphalism*. Baltimore, MD: Johns Hopkins University Press, 2008.

Designation Systems. "Boeing GQM-94 B-Gull." No Date. https://www.designation-systems.net/dusrm/m-94.html.

Designation Systems. "Bureau of Standards SWOD MK 9/ASM-N-2 Bat." No Date. https://www.designation-systems.net/dusrm/app1/asm-n-2.html.

Designation Systems. "E-Systems GQM-93." No Date. https://www.designation-systems.net/dusrm/m-93.html.

Designation Systems. "General Atomics RQ/MQ-1 Predator." No Date. https://www.designation-systems.net/dusrm/app2/q-1.html.

Designation Systems. "Lockheed MQM-105 Aquila." No Date. https://www.designation-systems.net/dusrm/m-105.html.

Designation Systems. "Martin SSM-A-1/B-61/TM-61/MGM-1 Matador." No Date. https://www.designation-systems.net/dusrm/m-1.html.

Designation Systems. "Martin TM-76/MGM-13/CGM-13 Mace." No Date. https://www.designation-systems.net/dusrm/m-13.html.

Designation Systems. "Matador." No Date. https://nuke.fas.org/guide/usa/theater/matador.htm.

Designation Systems. “Northrop AGM/BGM-136 Tacit Rainbow.” No Date. https://www.designation-systems.net/dusrm/m-136.html.

Designation Systems. “Raytheon (Texas Instruments) AGM-88 HARM.” No Date. https://www.designation-systems.net/dusrm/m-88.html.

Designation Systems. “Teledyne Ryan GQM-98 R-Tern.” No Date. https://www.designation-systems.net/dusrm/m-98.html.

Designation Systems. “Teledyne Ryan Q-2/KDA/xQM-34/BGM-34 Firebee.” No Date. https://www.designation-systems.net/dusrm/m-34.html.

Devore, Marc R. “Arms Production in the Global Village: Options for Adapting to Defense-Industrial Globalization.” *Security Studies* 22, No. 3 (2013): 532-72.

Ding, Jeffrey, and Allan Dafoe. “Engines of Power: Electricity, AI, and General-Purpose Military Transformations.” *European Journal of International Security* 8:3 (2023): 377-394.

Ding, Jeffrey, and Allan Dafoe. “The Logic of Strategic Assets: From Oil to AI.” *Security Studies* 30:2 (2021): 182-212.

Dombrowski, Peter and Eugene Gholz. *Buying Military Transformation: Technological Innovation and the Defense Industry*. New York: Columbia University Press, 2006.

Downes, Alexander B. “How Smart and Tough Are Democracies? Reassessing Theories of Democratic Victory in War.” *International Security* 33, No. 4 (2009): 9-51.

Downes, Alexander B. *Targeting Civilians in War*. Ithaca, NY: Cornell University Press, 2012.

Doyle, Michael W. “Kant, Liberal Legacies, and Foreign Affairs, Part 2.” *Philosophy and Public Affairs* 12, No. 3 (1983): 323-353.

Doyle, Michael W. “Kant, Liberal Legacies, and Foreign Affairs.” *Philosophy and Public Affairs* 12, No. 3 (1983): 205-235.

Drezner, Daniel. “Technological Change and International Relations.” *International Relations* 33, No. 2 (2019): 286-303.

Drone Wars. “Drone Crash Database.” No Date. https://dronewars.net/drone-crash-database/

Dungan, T.D. *V-2: A Combat History of the First Ballistic Missile*. Yardley, PA: Westholme Publishing, 2019.

Dunn, Rob. “The Writer Who Built the World’s First Engine-Powered Submarine.” Smithsonian Magazine, September 17, 2013. https://www.smithsonianmag.com/history/the-writer-who-built-the-worlds-first-engine-powered-submarine-8718089/.

Dupuy, Trevor N. *The Evolution of Weapons and Warfare*. Boston, MA: Da Capo Press, 1984.

Edlow, Sabrina R. U.S. Employment of Naval Mines: A Chronology. Alexandria, VA: Center for Naval Analyses, 1997, https://apps.dtic.mil/sti/tr/pdf/ADA362483.pdf.

Ehrhard, Thomas P. *Air Force UAVs: The Secret History*. Washington, D.C.: Mitchell Institute Press, 2010.

Eliason, Leslie C. and Emily O. Goldman, Editors, *The Diffusion of Military Technology and Ideas*. Palo Alto; CA: Stanford University Press, 2003.

Elman, Colin. "The Logic of Emulation: The Diffusion of Military Practices in the International System." Ph.D. Dissertation, Columbia University, 1999.

Evangelista, Matthew. *Innovation and the Arms Race: How the United States and the Soviet Union Develop New Military Technologies.* Ithaca, NY: Cornell University Press, 1988.

Evans, David and Mark Peattie. *Kaigun: Strategy, Tactics, and Technology in the Imperial Japanese Navy, 1887-1941*. Annapolis, MD: Naval Institute Press, 1997.

Evans, Stephen G., and Keith B. Delany. "The V1 (Flying Bomb) Attack on London (1944-1945); the Applied Geography of Early Cruise Missile Accuracy." *Applied Geography* 99 (2018): 44-53.

Everett, H.R. *Unmanned Systems of World Wars I and II*. Cambridge, MA: MIT Press, 2015.

Everstine, Brian W. "Elegy for the Predator." Air & Space Forces Magazine. May 29, 2018. https://www.airandspaceforces.com/article/elegy-for-the-predator/.

Farquhar, John T. *A Need to Know: The Role of Air Reconnaissance in War Planning, 1945-1953*. Maxwell Air Force Base, AL: Air University Press.

Feaver, Peter D. "Civil-Military Relations and Policy: A Sampling of a New Wave of Scholarship." *Journal of Strategic Studies* 40, No. 1-2 (2017): 325-342.

Feaver, Peter D. "Civil-Military Relations." *Annual Review of Political Science* 2 (1999): 211-241.

Federation of American Scientists, "AGM-136 Tacit Rainbow." No Date. https://man.fas.org/dod-101/sys/smart/agm-136.htm.

Federation of American Scientists, "M712 Copperhead." No Date. https://man.fas.org/dod-101/sys/land/m712.htm.

Federation of American Scientists. "AGM-88 HARM (High Speed Antiradiation Missile). No Date. http://man.fas.org/dod-101/sys/smart/agm-88.htm

Federation of American Scientists. "CL-289 / AN/USD-502." No Date. https://man.fas.org/dod-101/sys/ac/row/cl-289.htm.

Federation of American Scientists. "Mace." No Date. https://nuke.fas.org/guide/usa/theater/mace.htm.

Federation of American Scientists. "Pioneer Short Range Drone." No Date. https://irp.fas.org/program/collect/pioneer.htm.

Federation of American Scientists. “Senior Year / Aquatone / U-2 / TR-1.” No Date. https://irp.fas.org/program/collect/u-2.htm.

Finkel, Meir. *On Flexibility: Recovery from Technological and Doctrine Surprise on the Battlefield*. Stanford, CA: Stanford University Press, 2011.

Fino, Steven A. 2017. *Tiger Check: Automating the US Air Force Fighter Pilot in Air-to-Air Combat, 1950–1980*. Baltimore, MD: Johns Hopkins University Press, 2017.

Friedman, Norman. Fighters Over the Fleet: Naval Air Defense from Biplanes to the Cold War. Annapolis, MD: Naval Institute Press, 2016.

Friedman, Norman. *U.S. Submarines Through 1945: An Illustrated Design History*. Annapolis, MD: Naval Institute Press, 1995.

Fuhrmann, Matthew and Michael Horowitz. “Droning On: Explaining the Proliferation of Unmanned Aerial Vehicles.” *International Organization* 71, No. 2 (2017): 397–418.

Gady, Franz-Stefan. “Useful, but Not Decisive: UAVs in Libya’s Civil War.” International Institute for Strategic Studies. November 22, 2019. https://www.iiss.org/blogs/analysis/2019/11/mide-uavs-in-libyas-civil-war.

Galati, Gaspare. *100 Years of Radar*. New York, NY: Springer, 2015.

Garfinkel, Ben and Allan Dafoe. “How Does the Offense-Defense Balance Scale?” *Journal of Strategic Studies* 42, No. 6 (2019): 864-887.

Gartzke, Erik, and Jon R. Lindsay, “Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace.” *Security Studies* 24, No. 2 (2015): 316-348.

Gartzke, Erik and Yonatan Lupu. “Trading on Perceptions: Why World War I Was Not a Failure of Economic Interdependence.” *International Security* 36:4 (2012): 115-150.

Gettinger, Dan. “Drones Operating in Syria and Iraq.” Annandale-on-Hudson, N.Y.: Center for the Study of the Drone, Bard College, December 2016. https://dronecenter.bard.edu/files/2016/12/Drones-in-Iraq-and-Syria-CSD.pdf.

Gettinger, Dan. “The Drone Databook.” Annandale-on-Hudson, NY: Center for the Study of the Drone, Bard College, 2019., https://dronecenter.bard.edu/projects/drone-proliferation/databook/

Gholz, Eugene. “Globalization, Systems Integration, and the Future of Great Power War.” *Security Studies* 16, No. 4 (2007): 615-636.

Gilli, Andrea and Mauro Gilli. “The Diffusion of Drone Warfare? Industrial, Organizational, and Infrastructural Constraints.” *Security Studies* 25, No. 1 (2016): 50-84.

Gilpin, Robert. *War and Change in World Politics.* Cambridge, UK: Cambridge University Press, 1981.

Glaser, Charles L. “Realists as Optimists: Cooperation as Self Help.” *International Security* 19, No. 3 (1994/1995): 50-90.

Glaser, Charles L. “The Security Dilemma Revisited,” *World Politics* 50, No. 1 (1997): 171- 201.

Glaser, Charles L. “When Are Arms Races Dangerous? Rational Versus Suboptimal Arming.” *International Security* 28, No. 4 (2007): 44-84.

Glaser, Charles L., and Chaim Kaufmann, “What is the Offense-Defense Balance and Can We Measure It?” *International Security* 22, No. 4 (1998): 44-82.

Glaser, Charles L. *Rational Theory of International Politics*. Princeton, NJ: Princeton University Press, 2011.

Goldman, Emily O. “Cultural Foundations of Military Diffusion,” *Review of International Studies* 32, No. 1. (2006): 69–91.

Goldman, Emily O., and Richard B. Andres. “Systemic Effects of Military Innovation and Diffusion.” *Security Studies* 8, No. 4 (1999): 79-125.

Grant, Rebecca, *The Radar Game: Understanding Stealth and Aircraft Survivability.* (Arlington, VA: Mitchell Institute for Aerospace Studies, 2010).

Grauer, Ryan, and Michael Horowitz, “What Determines Military Victory? Testing the Modern System,” Security Studies 21, No. 1 (2012): 83–112.

Grauer, Ryan. *Commanding Military Power: Organizing for Victory and Defeat on the Battlefield*. Cambridge, UK: Cambridge University Press, 2016.

Gray, Edwyn. *The Devil’s Device: The Story of Robert Whitehead, Inventor of the Torpedo.* Annapolis, MD: Naval Institute Press, 1991.

Griffin, Stuart. “Military Innovation Studies: Multidisciplinary or Lacking Discipline?” *Journal of Strategic Studies* 40, No. 1–2 (2017): 196–224.

Grissom, Adam. “The Future of Military Innovation Studies.” *Journal of Strategic Studies* 29, No. 5 (2006): 905–34.

Gunitsky, Seva, *Aftershocks: Great Powers and Domestic Reforms in the Twentieth Century*. Princeton, NJ: Princeton University Press, 2017.

Gusterson, Hugh. *Drone: Remote Control Warfare.* Cambridge, MA: MIT Press, 2016.

Haas, Mark L. “A Geriatric Peace? The Future of U.S. Power in a World of Aging Populations.” *International Security* 32, No. 1 (2007):112-147.

Hacker, Barton C. “Military Institutions, Weapons, and Social Change: Toward a New History of Military Technology,” *Technology and Culture* 35, No. 4 (1994): 425-440.

Haider, Andre. *Remotely Piloted Aircraft Systems in Contested Environments: A Vulnerability Analysis.* Kalkar, Germany: Joint Air Power Competence Centre, 2014.

Hallion, Richard P. “A Troubling Past: Air Force Fighter Acquisition Since 1945.” *Airpower Journal* 4, No. 4 (1990): 4-23.

Hannah, Craig C. *Striving for Air Superiority: The Tactical Air Command in Vietnam*. College Station, TX.: Texas A&M University Press, 2002.

Hasik, James M. The Precision Revolution: GPS and the Future of Aerial Warfare. (Annapolis, MD: Naval Institute Press, 2002).

Haulman, Daniel L. “Manned Aircraft Losses Over the Former Yugoslavia, 1994-1999.” Air Force Historical Research Agency, 2015. https://www.afhra.af.mil/Portals/16/documents/Airmen-at-War/Haulman-MannedAircraftLossesYugoslavia1994-1999.pdf?ver=2016-08-22-131404-383.

Headrick, Daniel R. *The Invisible Weapon: Telecommunications and International Politics, 1851-1945*. Oxford: UK: Oxford University Press, 1991.

Helgason, Gudmunur. “The Walter U-boats.” Uboat.net. No Date. https://uboat.net/types/walter_hist.htm.

Helgason, Gudmunur. “Type UB I.” Uboat.net. no date. https://uboat.net/wwi/types/?type=UB+I.

Helgason, Gudmunur. “Type UB III.” Uboat.net. no date. https://uboat.net/wwi/types/?type=UB+III.

Helgason, Gudmunur. “Type UC I.” Uboat.net. no date. https://uboat.net/wwi/types/?type=UC+I.

Helgason, Gudmunur. “Type UC II.” Uboat.net. no date. https://uboat.net/wwi/types/?type=UC+II.

Helgason, Gudmunur. “Type UC III.” Uboat.net. no date. https://uboat.net/wwi/types/?type=UC+III.

Helgason, Gudmunur. “Type UC III.” Uboat.net. no date. https://uboat.net/wwi/types/?type=UC+III.

Herz, John H. “Idealist Internationalism and the Security Dilemma.” *World Politics* 2, No. 2 (1950): 157-180.

Hoffman, Frank G. 2021. *Mars Adapting: Military Change During War*. Annapolis, MD: Naval Institute Press. 2021.

Horowitz, Michael C. “Do Emerging Military Technologies Matter for International Politics?” *Annual Review of Political Science* 23 (2020): 385-400.

Horowitz, Michael C. and Shira Pindyck. “What is a Military Innovation and Why It Matters.” *Journal of Strategic Studies* 46, No. 1 (2023): 85-114.

Horowitz, Michael C. *The Diffusion of Military Power: Causes and Consequences for International Politics* (Princeton, NJ: Princeton University Press, 2010).

Horowitz, Michael, Joshua Schwartz, and Matthew Fuhrmann. “Who’s Prone to Drone? A Global Time-Series Analysis of Armed Uninhabited Aerial Vehicle Proliferation.” *Conflict Management and Peace Science* 39 No. 2 (2022): 119–142.

Horowitz, Michael, Sarah E. Kreps, and Matthew Fuhrmann. “Separating Fact from Fiction in the Debate over Drone Proliferation.” *International Security* 41, No. 2 (2016): 7-42.

Horowitz, Michael. “Artificial Intelligence, International Competition, and the Balance of Power.” *Texas National Security Review* 1, No. 3 (2018): 36-57.

Hunzeker, Michael A. *Dying to Learn: Wartime Lessons from the Western Front.* Ithaca, NY: Cornell University Press, 2021.

Imperial War Museum, “A Brief History of Drones.” No Date. https://www.iwm.org.uk/history/a-brief-history-of drones#:~:text=The%20Aerial%20Target%2C%20a%20British,was%20tested%20in%20March%201917.

Insinna, Valerie. “Air Force’s RQ-4 Global Hawk drones headed for retirement in FY27.” *Breaking Defense*. July 27, 2022. https://breakingdefense.com/2022/07/exclusive-air-forces-rq-4-global-hawk-drones-headed-for-the-boneyard-in-fy27/

Jackson, Frederick E. “Tannenberg: The First Use of Signals Intelligence in Modern Warfare.” *Strategic Research Project*. Carlisle, PA: US Army War College, 2002.

Jensen, Benjamin. *Forging the Sword: Doctrinal Change in the U.S. Army*. Stanford, CA: Stanford University Press, 2016.

Jervis, Robert. “Cooperation under the Security Dilemma.” *World Politics* 30, No. 2 (1978): 167–214.

Jervis, Robert. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University, 1976.

Jervis, Robert. *The Meaning of the Nuclear Revolution*. Ithaca, NY: Cornell University Press, 1989.

Jones, Marcus O. “Innovation for its Own Sake: The Type XXI U-boat.” *Naval War College Review* 67, No. 2 (2014): 1-14.

Jones, Mark C. “Give Credit Where Credit Is Due: The Dutch Role in the Development and Deployment of the Submarine Schnorkel.” *The Journal of Military History* 69, No. 4 (2005): 987-1012.

Jungdahl, Adam M. and Julia M. Macdonald, “Innovation Inhibitors in War: Overcoming Obstacles in the Pursuit of Military Effectiveness.” *Journal of Strategic Studies* 38, No. 4 (2015): 467-499.

Kaplan, Edward. *To Kill Nations: American Strategy in the Air-Atomic Age and the Rise of Mutually Assured Destruction*. Ithaca. NY: Cornell University Press, 2015.

Katzenstein, Peter J. *The Culture of National Security: Norms and Identity in World Politics*. New York: Columbia University Press, 1996.

Keaney, Thomas A., and Eliot A. Cohen. *Gulf War Air Power Survey: Summary Report.* Washington, DC: U.S. Government Printing Office, 1993.

Kennedy, Paul M. *The Rise and Fall of British Naval Mastery*. London, UK: Allen Lane, 1976.

Kier, Elizabeth. *Imagining War: French and British Military Doctrine Between the Wars.* Princeton, NJ: Princeton University Press, 1997.

Knight, Roger. "British North Atlantic Convoys, 1812–14, and the Subsequent Rejection of the Convoy System." in *Navies in Multipolar Worlds: From the Age of Sail to the Present*. edited by Paul Kennedy and Evan Wilson. London, UK: Routledge, 2021.

Kollars, Nina. "Military Innovation's Dialectic: Gun Trucks and Rapid Acquisition." *Security Studies* 23, No. 4: 787–813.

Kontorovich, Vladimir. *The Impoverished Superpower: Perestroika and the Burden of Soviet Military Spending.* San Fransisco, CA: Institute for Contemporary Studies Press, 1990.

Kopp, Carlo. "The Dawn of the Start Bomb." Air Power Australia. March 26, 2011. https://www.ausairpower.net/WW2-PGMs.html.

Kreis, John F. "Unmanned Aircraft in Israeli Air Operations." *Air Power History* 37, No. 4 (1990): 46-50.

Kreps, Sarah and Micah Zenko. "The Next Drone Wars: Preparing for Proliferation." *Foreign Affairs* 93, No. 2 (2014): 68–79.

Kreps, Sarah E. *Drones: What Everyone Needs to Know.* Oxford, UK: Oxford University Press, 2016.

Kreuzer, Michael P. *Drones and the Future of Air Warfare: The Evolution of Remotely Piloted Aircraft*. London, UK: Routledge, 2016.

Kuo, Kendrick. "Military Magic: The Promise and Peril of Military Innovation." Ph.D. dissertation, Political Science, The George Washington University, 2021.

Lagrone, Sam. "Mabus: F-35 Will Be 'Last Manned Strike Fighter' the Navy, Marines 'Will Ever Buy or Fly'." *USNI News*. April 15, 2015. https://news.usni.org/2015/04/15/mabus-f-35c-will-be-last-manned-strike-fighter-the-navy-marines-will-ever-buy-or-fly.

Lake, David A. "Escape from the State of Nature: Authority and Hierarchy in World Politics." *International Security* 32, No. 1 (2007): 47-79.

Lake, David A. "Powerful Pacifists: Democratic States and War." American Political Science Review 80, No. 1 (1992): 24-37.

Lambeth, Benjamin S. "Kosovo and the Continuing SEAD Challenge." *Aerospace Power Journal* 16, No. 2 (2002): 8-22.

Lambeth, Benjamin S., *NATO's Air War for Kosovo: A Strategic and Operational Assessment.* Santa Monica, CA.: RAND, 2001.

Lautenschlager, Karl. "Technology and the Evolution of Naval Warfare." *International Security* 8, No. 2 (1983): 3-51.

Lautenschlager, Karl. "The Submarine in Naval Warfare, 1901-2001." *International Security* 11, No. 3 (1986): 94-140.

Levy, Jack S. "The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis." *International Studies Quarterly* 28, No. 2 (1984): 219-238.

Lieber, Keir A., "Grasping the Technological Peace: The Offense-Defense Balance and International Security." *International Security* 25, No. 1 (2000): 71-104.

Lieber, Kier A. *War and the Engineers: The Primacy of Politics Over Technology*. Ithaca, NY: Cornell University Press, 2005.

Lindsay, Jon R. "Demystifying the Quantum Threat: Infrastructure, Institutions, and Intelligence Advantage." *Security Studies* 29, No. 2 (2020): 335–61.

Lindsay, Jon R. "Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence against Cyber Attack." *Journal of Cybersecurity* 1, No. 1 (2015): 53–67.

Lindsay, Jon R. *Information Technology and Military Power*. Ithaca, NY: Cornell University Press, 2020.

Losey, Stephen. "US Air Force eyes fleet of 1,000 drone wingmen as planning accelerates." *Defense News*. March 8, 2023. https://www.defensenews.com/air/2023/03/08/us-air-force-eyes-fleet-of-1000-drone-wingmen-as-planning-accelerates/.

Lowell, Bernard. *Echoes of War: The Story of H2S Radar*. Boca Raton, FL: CRC Press, 1991.

Lyall, Jason. *Divided Armies: Inequality and Battlefield Performance in Modern War*. Princeton, NJ: Princeton University Press, 2020.

Lynn-Jones, Sean M. "Offense-Defense Theory and Its Critics." *Security Studies* 4, No. 4 (1995): 660-694.

MacKenzie, *Donald. Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance*. Cambridge, MA: MIT Press, 1990.

Macksey, Kenneth. *Technology in War: The Impact of Science on Weapon Development and Modern Battle.* New York: Simon and Shuster, 1986.

Mahnken, Thomas. *Uncovering Ways of War: U.S. Intelligence and Foreign Military Innovation, 1918-1941*. Ithaca, NY: Cornell University Press, 2002.

Mallmann-Showell, Jak P. *German Naval Codebreakers*. Annapolis, MD: Naval Institute Press, 2003.

Marcus, Raphael. "Military Innovation and Tactical Adaptation in the Israel–Hezbollah Conflict: The Institutionalization of Lesson-Learning in the IDF." *Journal of Strategic Studies* 38, No. 4 (2015): 344-370.

Mayer, Michael. "The New Killer Drones: Understanding the Strategic Implications of Next Generation Unmanned Combat Aerial Vehicles." *International Affairs* 91, No. 4 (2015): 765-780.

McCarthy, Daniel R. *Power, Information Technology, and International Relations Theory: The Power and Politics of US Foreign Policy and Internet.* London, UK: Palgrave Macmillan, 2015.

McNeil, William H. *The Pursuit of Power: Technology, Armed Force, and Society Since A.D. 1000*. Chicago, IL: University of Chicago Press, 1982.

Mearsheimer, John J. *The Tragedy of Great Power Politics*. New York: Norton, 2001.

Messimer, Dwight R. *Find and Destroy: Antisubmarine Warfare in World War I*. Annapolis: MD: Naval Institute Press, 2001.

Michel III, Marshall L. *Clashes: Air Combat over North Vietnam 1965–1972*. Annapolis, MD: Naval Institute Press, 1997.

Miller, Edward S. *War Plan Orange: The U.S. Strategy to Defeat Japan, 1897-1945*. Annapolis, MD: Naval Institute Press, 1991.

Miller, Steve. *The 99th Strategic Reconnaissance Squadron: The Air Force's Story of Unmanned Reconnaissance in the Vietnam War*. Miller Management Systems, 2017.

Millett, Allan R. and Williamson Murray, Editors., *Military Effectiveness*, Vols. 1–3. Boston: Allen and Unwin, 1988.

Millett, Allan R., Williamson Murray, and Kenneth H. Watman, "The Effectiveness of Military Organizations." *International Security* 11, No. 1 (1986): 37-71.

Mir, Asfandyar and Dylan Moore. "Drones, Surveillance, and Violence: Theory and Evidence from a US Drone Program." *International Studies Quarterly* 63 No. 4 (2019): 846-862.

Mir, Asfandyar, "What Explains Counterterrorism Effectiveness? Evidence from the US Drone War in Pakistan," *International Security* 43, No. 2 (2018): 45-83.

Mitzen, Jennifer. "Ontological Security in World Politics: State Identity and the Security Dilemma." *European Journal of International Relations*, Vol. 12, No. 3 (2006): 341-370.

Mueller, John. "The Perfect Enemy: Assessing the Gulf War," *Security Studies* 5, No. 1 (1995): 77-117.

Mukunda, Gautam, "We cannot go on: Disruptive innovation and the First World War Royal Navy." *Security Studies* 19, No. 1 (2010): 124-159.

Murray, Williamson, and Allan R. Millett, Editors, *Military Innovation in the Interwar Period*. Cambridge, UK: Cambridge University Press, 2009.

Musee National De La Marine. "The Model of Le Plongeur." No date. https://www.musee-marine.fr/en/collections/rochefort/the-model-of-le-plongeur.html.

National Air and Space Museum. "Gyroscope, Autopilot, Kettering Liberty Eagle 'Bug'." No Date. https://airandspace.si.edu/collection-objects/gyroscope-autopilot-kettering-liberty-eagle-bug/nasm_A20010524000.

National Air and Space Museum. "Pioneer RQ_2A UAV." No Date. https://airandspace.si.edu/collection-objects/pioneer-rq-2a-uav/nasm_A20000794000.

National Museum of the U.S. Navy. "Hedgehog Weapons." No Date. https://www.history.navy.mil/content/history/museums/nmusn/explore/photograp

hy/wwii/wwii-atlantic/battle-of-the-atlantic/anti-submarine-warfare/hedgehog.html.

National Museum of the U.S. Navy. "Mousetrap Weapons." No Date. https://www.history.navy.mil/content/history/museums/nmusn/explore/photography/wwii/wwii-atlantic/battle-of-the-atlantic/anti-submarine-warfare/mousetrap.html.

National Museum of the United States Air Force, "Kettering Aerial Torpedo 'Bug'." No Date. https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/198095/kettering-aerial-torpedo-bug/.

National Museum of the United States Air Force. "Lockheed D-21B." No Date. https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/195778/lockheed-d-21b/.

National Museum of the United States Air Force. "Republic/Ford JB-2 Loon (V-1 Buzz Bomb)." No Date. https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/196227/republicford-jb-2-loon-v-1-buzz-bomb/.

Naval History and Heritage Command. "H.L. Hunley Wreck (1864)." no date, https://www.history.navy.mil/research/underwater-archaeology/sites-and-projects/ship-wrecksites/hl-hunley.html.

Naval History and Heritage Command. "NH 2058 Drawing of Robert Fulton's NAUTILUS, 1798." no date, https://www.history.navy.mil/content/history/nhhc/our-collections/photography/numerical-list-of-images/nhhc-series/nh-series/NH-02000/NH-2058.html.

Naval History and Heritage Command. "The Submarine Turtle: Naval Documents of the Revolutionary War." November 2, 2017. https://www.history.navy.mil/research/library/online-reading-room/title-list-alphabetically/s/submarine-turtle-naval-documents.html.

Nedelchev, Nickolay. "The Norwegian Operation and the Torpedo Crisis." Uboat.net. April 2, 1999. https://uboat.net/articles/23.html.

Neufeld, Michael. *The Rocket and the Reich: Peenemunde and the Coming of the Ballistic Missile Era*. Cambridge, MA: Harvard University Press, 1996.

O'Connor, Sean. "Arming America's Interceptors: The Hughes Falcon Missile Family." Air Power Australia. June 2011. https://www.ausairpower.net/Falcon-Evolution.html.

Paarlberg, Robert. "Knowledge as Power: Science, Military Dominance, and U.S. Security." *International Security* 29:1 (2004): 122-151.

Pape, Robert A. *Bombing to Win: Air War and Coercion in War*. Ithaca, NY: Cornell University Press, 1996.

Pavelec, Sterling Michael. *The Jet Race and the Second World War*. Westport, CT: Prager, 2007.

Pedlow, Gregory, Donald Welzenbach, and Chris Pockock. *The Central Intelligence Agency and Overhead Reconnaissance: The U-2 and OXCART Programs, 1954-1974*. New York: Skyhorse, 2016.

Pike, Jonathan. "Tu-123." Global Security. No Date, https://www.globalsecurity.org/military/world/russia/tu-123.htm.

Pike, Jonathan. "Tu-141 Strizh / Tu-143 "Reis." Global Security. No Date, https://www.globalsecurity.org/military/world/ukraine/tu-141-ssm.htm.

Pike, Jonathan. "Tu-143 Reys." Global Security. No Date, https://www.globalsecurity.org/military/world/russia/tu-143.htm.

Pollack, Kenneth M. "The Influence of Culture on Arab Military Effectiveness." Ph.D. dissertation, Massachusetts Institute of Technology, 1996.

Pollack, Kenneth. *Arabs at War*. Lincoln, NE: University of Nebraska Press, 2002.

Polmar, Norman. "Historic Aircraft – The Pioneering Pioneer." *Naval History Magazine* 27, No. 5 (2013) https://www.usni.org/magazines/naval-history-magazine/2013/september/historic-aircraft-pioneering-pioneer.

Polmar, Norman. "Historical Aircraft – Silent (by) Night." *Naval History Magazine* 31, No. 5 (2017), https://www.usni.org/magazines/naval-history-magazine/2017/october/historic-aircraft-silent-night.

Posen, Barry R. "Command of the Commons: The Military Foundation of U.S. Hegemony." *International Security* 28, No. 1 (2003): 5–46.

Posen, Barry R. "Nationalism, the Mass Army, and Military Power." *International Security* 18, No. 2 (Fall 1993): 80-124.

Posen, Barry R. "The Security Dilemma and Ethnic Conflict." *Survival* 35, No. 1 (1993): 27-47.

Posen, Barry R. "The War for Kosovo: Serbia's Political-Military Strategy." *International Security* 24, No. 4 (2000), pp. 39–84.

Posen, Barry R. *The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars.* Ithaca, NY: Cornell University Press, 1984.

Posen, Stephen P. *Winning the Next War: Innovation and the Modern Military*. Ithaca, NY: Cornell University Press, 1991.

Press, Daryl G. "Lessons from Ground Combat in the Gulf: The Impact of Training and Technology." International Security 22, No. 2 (1997), 137-146.

Press, Daryl G. "The Myth of Air Power in the Persian Gulf War and the Future of Warfare." *International Security* 26, No. 2 (2001): 5–44.

Public Broadcasting Service. "Remote Control." No Date. https://www.pbs.org/tesla/ins/lab_remotec.html.

Quester, George. *Offense and Defense in the International System*. New York: John Wiley Publishers, 1977.

Quinlivan, James. “Coup-Proofing: Its Practice and Consequences in the Middle East.” *International Security* 24, No. 2 (1999): 131–165.

Raudzens, George. “War-Winning Weapons: The Measurement of Technological Determinism in Military History.” *The Journal of Military History* 54, No. 4 (1990): 403-434.

Reim, Garrett. “Record Number of UAV Shoot Downs Prompt New USAF Tactics and Countermeasure Pod.” *Flight Global*. June 30, 2020. https://www.flightglobal.com/military-uavs/record-number-of-uav-shoot-downs-prompt-new-usaf-tactics-and-countermeasure-pod/138908.article.

Reiter, Dan and Allan C. Stam. “Democracy and Battlefield Military Effectiveness.” *Journal of Conflict Resolution* 42, No. 3 (1998): 259-277.

Reiter, Dan and Allan C. Stam. *Democracies at War*. Princeton, NJ: Princeton University Press, 2002.

Reiter, Dan and Allan Stam, “Democracy and Battlefield Military Effectiveness.” *Journal of Conflict Resolution* 42, No. 3 (1998): 377–89.

Reiter, Dan. “Nationalism and Military Effectiveness: Post-Meiji Japan.” in *Creating Military Power: The Sources of Military Effectiveness*. edited by Risa A. Brooks and Elizabeth A. Stanley. Stanford, CA: Stanford University Press, 2007.

Resende-Santos, Joao, “Anarchy and the Emulation of Military Systems: Military Organization and Technology in South America, 1870-1930.” *Security Studies* 5, No. 3 (1996): 193–260.

Riordan, Michael and Lillian Hoddeson. *Crystal Fire: The Birth of the Information Age*. New York: W.W. Norton & Company, 1997.

Røksund, Arne. *The Jeune Ecole: The Strategy of the Weak*. Leiden, Netherlands: Brill, 2007.

Roland, Alex. “Science, Technology, and War.” *Technology and Culture* 36, No. 2 (1995): 83-100.

Rosen, Stephen P. “Military Effectiveness: Why Society Matters.” *International Security* 19, No. 4 (1995): 5–31.

Rosen, Stephen P. *Societies and Military Power: India and Its Armies.* Ithaca, NY: Cornell University Press, 1996.

Ruttan, Vernon W. *Is War Necessary for Economic Growth? Military Procurement and Technology Development*. Oxford, UK: Oxford University Press, 2006.

Sagan, Scott. “1914 Revisited: Allies, Offense, and Instability.” *International Security* 11, No. 2 (1986): 151-175.

Sapolsky, Harvey M. *The Polaris System Development: Bureaucratic and Programmatic Success in Government*. Cambridge, MA: Harvard University Press, 1972.

Schneider, Jacquelyn. “The Capability/Vulnerability Paradox and Military Revolutions: Implications for Computing, Cyber, and the Onset of War.” *Journal of Strategic Studies* 42, No. 6 (2019): 841-863.

Schultz, Timothy P., The Problem with Pilots: How Physicians, Engineers, and Airpower Enthusiasts Redefined Flight. (Baltimore, MD: Johns Hopkins University Press, 2018).

Schweller, Randall L. "Neorealism's Status-Quo Bias: What Security Dilemma?" *Security Studies* 5, No. 3 (1996): 90-121.

Sechser, Todd S., Neil Narang, and Caitlin Talmadge, Editors, *Emerging Technologies and International Stability.* New York: Routledge, 2022.

Shah, Aqil. "Do U.S. Drone Strikes Cause Blowback? Evidence from Pakistan and Beyond." *International Security* 42, No. 4 (2018): 47-84.

Shanker, Thom. "Threats and Responses: Combat; Iraqi Planes Shoot Down American Predator Drone." *New York Times*. December 24, 2002. https://www.nytimes.com/2002/12/24/world/threats-and-responses-combat-iraqi-planes-shoot-down-american-predator-drone.html.

Singer, P.W. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. New York: The Penguin Press, 2009.

Slayton, Rebecca. "What is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment." *International Security* 41, No. 3 (2017): 72-109.

Smithsonian. "Bomb, Guided, Fritz X (X-1)." No Date. https://www.si.edu/object/bomb-guided-fritz-x-x-1%3Anasm_A19840794000.

Snyder, Jack. "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984." *International Security* 9, No. 1 (1984): 108-146.

Stein, Aaron, "Say Hello to Turkey's Little Friend: How Drones Help Level the Playing Field." War on the Rocks, June 11, 2021. https://warontherocks.com/2021/06/say-hello-to-turkeys-little-friend-how-drones-help-level-the-playing-field/.

Stillion, John, and Bryan Clark. *What it Takes to Win: Succeeding in 21st Century Battle Network Competitions.* Washington, DC: Center for Strategic and Budgetary Assessments, 2015.

Strachan, *The First World War in Africa*. Oxford University Press, 2004.

Stulberg, Adam. "Managing the Unmanned Revolution in the U.S. Air Force." *Orbis* 51, No. 2 (2007): 251–265.

Sullivan, Jeffrey M. "Evolution of Revolution? The Rise of UAVs." *IEEE Technology and Society Magazine* 25, No. 3 (2006): 43-49

Swords, S.S. *Technical History of the Beginnings of Radar*. Stevenage, UK: Institution of Engineering and Technology, 2008.

Syrett, David. *The Battle of the Atlantic and Signals Intelligence: U–Boat Situations and Trends, 1941–1945*. London, UK: Routledge, 2002.

Talmadge, Caitlin. "Different Threats, Different Militaries: Explaining Organizational Practices in Authoritarian Armies." *Security Studies* 25, No. 1 (2016): 111–41.

Talmadge, Caitlin. “Emerging Technology and Intra-War Escalation Risks: Evidence from the Cold War, Implications for Today.” *Journal of Strategic Studies* 42, No. 6 (2019): 864-887.

Talmadge, Caitlin. *The Dictator’s Army: Battlefield Effectiveness in Authoritarian Regimes.* Ithaca, NY: Cornell University Press, 2015.

Tang, Shiping. “The Security Dilemma: A Conceptual Analysis.” *Security Studies* 18, No. 3 (2009): 587-623.

The Economist. “How Cheap Drones Are Transforming War in Ukraine.” February 5, 2024. https://www.economist.com/interactive/science-and-technology/2024/02/05/cheap-racing-drones-offer-precision-warfare-at-scale.

The Economist. “Killer Drones Pioneered in Ukraine Are the Weapons of the Future.” February 8, 2024. https://www.economist.com/leaders/2024/02/08/killer-drones-pioneered-in-ukraine-are-the-weapons-of-the-future.

The Museum of Flight. “Lockheed YO-3A Quiet Star.” No Date. https://www.museumofflight.org/exhibits-and-events/aircraft/lockheed-yo-3a-quiet-star.

Trachtenberg, Marc. “Assessing Soviet Economic Performance During the Cold War: A Failure of Intelligence?” *Texas National Security Review* 1, No. 2 (2018): 76-101.

Tracy, Cameron L., and David Wright. “Modelling the Performance of Hypersonic Boost-Glide Missiles.” *Science & Global Security* 28, No. 3 (2020): 135-170.

Trevithick, Joseph. “Houthi Rebels Shoot Down U.S. Air Force MQ-9 Reaper Over Yemen.” The War Zone. June 29, 2019. https://www.twz.com/14806/houthi-rebels-shoot-down-u-s-air-force-mq-9-reaper-over-yemen.

van Creveld, Martin. *Technology and War: From 2000 B.C. to the Present*. New York: Free Press, 1989.

Van Evera, “Offense, Defense, and the Causes of War.” *International Security* 22, No. 4 (1998): 5-43.

Van Evera, Stephen. “The Cult of the Offensive and the Origins of the First World War.” *International Security* 9, No. 1 (1984): 58-107.

Van Evera, Stephen. *Causes of War: Power and the Roots of Conflict*. Ithaca, NY: Cornell University Press, 1999.

Van Valen, Leigh. “A New Evolutionary Law.” *Evolutionary Theory* 1, No. 2 (1973): 1-30.

Wagner, Abraham R. The Lessons of Modern War: Volume 1: The Arab-Israeli Conflicts, 1973–1989. Boulder, CO: Westview, 1990.

Walt, Stephen M. *The Origins of Alliances*. Ithaca, NY: Cornell University Press, 1990.

Wanger, William. *Lightning Bugs and Other Reconnaissance Drones.* Fallbrook, CA: Aero Publishers, 1982.

Watson Jr., Raymond, Radar Origins Worldwide: History of its Evolution in 13 Nations Through World War II. Bloomington, IN: Trafford Publishing, 2009.

Watts, Barry D. *Six Decades of Guided Munitions and Battle Networks: Progress and Prospects*. Washington, D.C.: Center for Strategic and Budgetary Assessments, 2007.

Weinberger, Seth. "Institutional Signaling and the Origin of the Cold War." *Security Studies* 12, No. 4 (2004): 80-115.

Wendt, Alexander. "Anarchy Is What States Make of It." *International Organization* 46, No. 2 (1992): 391-425.

Werrell, Kenneth P. Archie to SAM: A Short Operational History of Ground-Based Air Defense. Maxwell Air Force Base, AL: Air University Press, 2005).

Werrell, Kenneth P. Chasing the Silver Bullet: U.S. Air Force Weapons Development from Vietnam to Desert Storm. Washington, D.C.: Smithsonian Institution Scholarly Press, 2003.

Wildenberg, Thomas. "A Sub-Hunting Bloodhound." *Naval History Magazine* 31, No. 5 (2017).

Wildenberg, Thomas. "Origins of Inertial Navigation." *Air Power History* 63, No. 4 (2016): 17.24.

Wildenberg, Thoms and Norman Polmar. *Ship Killer: A History of the American Torpedo*. Annapolis, MD: Naval Institute Press, 2010.

Wright, Kevin. *We Were Never There: Volume 1: CIA U-2 Operations Over Europe, the USSR, and the Middle East, 1956-1960.* Warwick, UK: Helion and Company, 2021.

Wright, Kevin. *We Were Never There: Volume 2: CIA U-2 Asia and Worldwide Operations 1957-1974.* Warwick, UK: Helion and Company, 2021.

Wrigley, W. "History of Inertial Navigation." *Navigation* 24, No. 1 (1977): 1-6.

Yogev, Haim, Ronen A. Cohen, and Eyal Lewin. "Revolution in Military Affairs – The Operation Mole Cricket as a Case Study for the Technological Race During the Cold War." *International Area Studies Review* 25, No. 2 (2022): 138-156.

Zaloga, Steven J. "Defending the Capitals: The First Generation of Soviet Strategic Air Defense Systems 1950–1960." *Journal of Slavic Military Studies* 10, No. 4 (1997): 30-43.

Zaloga, Steven J. *Soviet Air Defence Missiles: Design, Development and Tactics*. London, UK: Jane's Information Group, 1989.

Zarzecki, Thomas Z. *Arms Diffusion: The Spread of Military Innovations in the International System.* London, UK: Routledge, 2002.

Zegart, Amy. "Cheap Fights, Credible Threats: The Future of Armed Drones and Coercion." *Journal of Strategic Studies* 43, No. 1 (2020): 6-46.

Zisk, Kimberly Marten. *Engaging the Enemy: Organization Theory and Soviet Military Innovation, 1955-1991*. Princeton, N.J.: Princeton University Press, 1993.