Taking Offense at Offense-Defense Theory
Author(s): James W. Davis, Jr., Bernard I. Finel, Stacie E. Goddard, Stephen Van Evera, Charles L. Glaser and Chaim Kaufmann
Reviewed work(s):
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/2539342

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
Correspondence
Taking Offense at Offense-Defense Theory

James W. Davis, Jr.
Bernard I. Finel
Stacie E. Goddard
Stephen Van Evera
Charles L. Glaser and
Chaim Kaufmann

To the Editors (James W. Davis, Jr. writes):

In his article “Offense, Defense, and the Causes of War,”1 Stephen Van Evera claims that “offense-defense theory” is “important,” has “wide explanatory range. . . . wide real-world applicability. . . . large prescriptive utility. . . . [and] is quite satisfying” (p. 41).

Van Evera’s conclusions are, however, unwarranted. First, his reformulation of influential arguments made prominent by Robert Jervis stretches the meaning of key concepts such that interesting avenues of empirical inquiry are closed off rather than opened. Second, the hypotheses—or “prime predictions”—Van Evera derives from the theory are themselves products of faulty deductive logic. Furthermore, they are non-testable, presumably nonscientific in Van Evera’s understanding of the term.2 Van Evera’s results are thus of little use to the social scientist who is interested in understanding the myriad causes of war and conditions facilitative of peace.

In his classic article, “Cooperation under the Security Dilemma,” Jervis argued that the security dilemma is more virulent and the international system less stable when offense enjoys an advantage over defense. By contrast, when defense is more potent, status quo powers find it easier to adopt compatible security policies, and the pernicious effects of international anarchy are greatly diminished.3 Although the operation-

James W. Davis, Jr., is Assistant Professor of International Politics at Ludwig-Maximilians-Universität in Munich, Germany. He is also a NATO Research Fellow.

Bernard I. Finel is Associate Director of the National Security Studies Program and Visiting Assistant Professor of National Security Studies and International Affairs at Georgetown University’s Edmund A. Walsh School of Foreign Service. He thanks Dieter Dette, Robert Hafna, Tim Hoyt, Jeffrey Lord, Kristin Lord, James Ludes, Gary Schaub, and Brent Sterling for their comments.

Stacie E. Goddard is a doctoral candidate in the Department of Political Science at Columbia University.

Stephen Van Evera teaches International Relations in the Political Science Department at the Massachusetts Institute of Technology. He is a member of the MIT Security Studies Program.

Charles L. Glaser is Professor and Deputy Dean in the Irving B. Harris Graduate School of Public Policy Studies at the University of Chicago. Chaim Kaufmann is Associate Professor of International Relations at Lehigh University.


© 1998 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.
alization of the offense-defense balance has been the subject of considerable debate,⁴ the concept as originally employed by Jervis referred to the modalities of battlefield conquest: military tactics, strategy, technology, and the state’s geography. The argument's appeal derives from its elegance and parsimony, as well as its explanatory range. Through variations in a rather simple—essentially material—relationship, we appear to gain leverage over a wide range of behavioral outcomes.

In his reformulation of the offense-defense balance, however, Van Evera adds “diplomatic factors” to the military and geographic factors identified by Jervis. In doing so, Van Evera subsumes under the offense-defense balance much of what we thought the balance helped explain. When “collective security systems, defensive alliances, and balancing behavior by neutral states” (pp. 21–22) are all constitutive of the offense-defense balance, we are no longer in a position to ask which military and geographic factors promote balancing, bandwagoning, or efforts at collective security: how they do so; or how the balance between offense and defense interacts with these diplomatic variables to produce such outcomes as war, peace, or overall system stability. Van Evera’s redefinition of the offense-defense balance is a step backward, a regressive reformulation of a heretofore useful concept.

A second problem emerges because Van Evera fails to keep the material or “objective” offense-defense balance analytically distinct from the balance as it is perceived by the actors. That actors might not apprehend the true or objective state of the offense-defense balance was already recognized by Jervis.⁵ The manifest difficulties that discrepancies between the objective and perceptual balance raise for attempts to use the concept in actual empirical investigations were, however, only later appreciated.⁶

The individual theorist may come down on one or the other side of the objective/perceptual divide, or she may choose to test which of the two variants accounts for outcomes in a given case. Logically ruled out, however, is the combination of both in a given hypothesis. Yet this is precisely what Van Evera attempts to do: “War will be more common in periods when conquest is easy or is believed easy, less common when conquest is difficult or is believed difficult” (p. 22). As formulated, the hypothesis is imprecise, internally incoherent, and as a result cannot be tested in any meaningful fashion.


If two dimensions are at work—one objective and one perceptual—then logically we have four possible combinations. That is, conquest can be (1) easy and believed to be easy; (2) easy but believed to be difficult; (3) difficult but believed to be easy; or (4) difficult and believed to be difficult.

At least two cases contemplated by the permutation of the two variants of the offense-defense balance cannot be included in the same hypothesis, because they stand in logical opposition to each other. Thus we cannot have a hypothesis that simultaneously predicts war to be relatively frequent because people mistakenly believe offense to be dominant and rare because the defense is in fact dominant. Similarly, we cannot have a situation where war is predicted to be rare because the defense is believed to be dominant, but where in fact offense is dominant and the hypothesis simultaneously predicts wars to be more frequent. And if perceptions always track the “objective” offense-defense balance, then parsimony would dictate we leave perceptions out of our theory and thus reject the two classes of cases emerging from the prime prediction that are not ruled out by logic (i.e., offense is dominant and believed to be dominant, and defense is dominant and believed to be dominant). Moreover, such cases would be uninformative if we are interested in finding out how perceptions matter.

Third, Van Evera overstates the extent to which his theory stands up to empirical tests. He argues that “the strength of a passed test depends on the uniqueness of the predictions tested. Do other theories predict the outcome observed, or is the prediction unique to the tested theory? The predictions here seem quite unique. There is no obvious competing explanation for the periodic upsurges and downsurges in European expansionism and warfare outlined above. Offense-defense theory has the field to itself” (p. 35).

Every hypothesis is, however, tested against a competing explanation, even if merely a hypothetical counterfactual. But given that every outcome is in some way consistent with Van Evera’s hypothesis, one cannot even formulate a hypothetical counterfactual. Moreover, it is generally accepted that one is justified in ascribing some plausibility to a theory’s explanatory claims only after it has been tested against a competing theory. Theories and hypotheses are “fortified” or “strengthened” to the degree to which they pass tests that are suggested in light of competing explanations. Van Evera is thus

8. This has the effect of closing off a traditional escape route for structural theorists (i.e., the argument that their theory explains only tendencies and not particular outcomes, because tendencies are demonstrable only to the extent to which we can clearly identify outliers).
10. For discussions of hard or crucial tests in the social sciences, see Harry Eckstein, “Case Study and Theory in Political Science,” in Fred Greenstein and Nelson Polsby, eds., Handbook of Political...
promoting an unorthodox understanding of hard tests when he writes: “Alternative explanations for the rise and fall of American global activism are hard to come up with, leaving the offense-defense theory's explanation without strong competitors, so this element of the test posed by the U.S. case is fairly strong” (p. 40).

Of course, serious alternative explanations for variations in war propensity abound. For example, based on a study of the European states system from 1640 to 1990, Andreas Osiander concluded that stability is a function of the coherence of the principal (normative) assumptions upon which an international system is founded. And although he does not dismiss the effects of “size, structure, power, and geographic position of the various European states,” Paul Schroeder argues that the chief difference between the relatively war-prone late eighteenth century and the more peaceful Concert of Europe a generation later was the lack of consensus among the great powers on legitimate principles of conduct and an equitable balance of power prior to the Napoleonic Wars.

Given the existence of competitors, Van Evera’s discovery that he has the field to himself suggests that he is either lost, or is playing something more akin to solitaire than to science.

—James W. Davis, Jr. 
Munich, Germany

To the Editors (Bernard I. Finel writes):

Several recent articles have provided textured considerations of offense-defense theory and the impact of the offense-defense balance on state behavior. These works have tightened the conceptual logic and added much-needed refinements to the argument. Four major problems with offense-defense theory remain, however. First, offense-defense theory ignores interaction effects in warfare. Second, it makes ill-considered assumptions about the links between control of territory, conquest, and victory in war. Third, the theory is still neither well conceptualized nor operationalized. Finally, the approach lacks parsimony.

The Importance of Interaction Effects

The offense-defense balance is not a structural variable. Rather it can be influenced by immediate decisions about deployments and employment strategies. This fact creates difficulties for Charles Glaser and Chaim Kaufmann as they try to use the tools of net assessment to operationalize and potentially quantify the offense-defense balance. Their article is vague about the meaning of net assessment. Are Glaser and Kaufmann referring to net assessment as used by Eliot Cohen, Andrew Krepinevich, and Andrew Marshall—that is, broad-based, subjective analyses of nonmilitary as well as military factors? Or are they referring to campaign analysis—that is, the use of mathematical models to predict the results of highly specified force-on-force engagements? Campaign analysis would certainly fit their goal, but the problem is that campaign analysis usually relies on Forward Edge of the Battle Area (FEBA) models. FEBA models are useful in explaining the results of attrition warfare, but not necessarily of dynamic, maneuver-based warfare.

To predict the outcome of dynamic, maneuver-based warfare, it is possible to use complex war games. These war games rarely create reproducible results, and they are extremely sensitive to modification in the initial rules. However, war games usually demonstrate that different strategies, and more important, the interaction effects of different strategies, make a big difference. If we take the war-game approach seriously, then we must conclude that the offense-defense balance is not a structural variable, but an outgrowth of strategic interaction.

The effectiveness of battlesystems depends on employment strategies, doctrine, and training and tactics. Changes at these three distinct levels of analysis are potentially independently capable of altering the course of a battle. For example, the Schlieffen Plan determined the initial course of World War I in the West at the strategic level. It determined how the Germans would mobilize their forces, their concentration points, and their operational goals. In the end, the Schlieffen Plan’s flaws—overextension and an uncovered right flank—doomed the German attack. Doctrine refers to the conceptual

3. Ibid., p. 76.
6. As an empirical example, the German attack on France in 1940 did not succeed quickly because of offensive dominance. Rather the German advance through the Ardennes to the English Channel coast was particularly effective because the Franco-British forces were pivoting into Belgium at the time. Martin van Creveld et al., Air Power and Maneuver Warfare (Maxwell Air Force Base, Ala.: Air University Press, 1994), p. 41.
basis for a tactical battlesystem. The blitzkrieg doctrine was a complex melding of armor, airpower, and disruptive penetrating advances. The development of the blitzkrieg concept allowed for the exploitation of the emerging technologies of the pre-World War II period, thus leading to a discontinuous increase in military effectiveness. Training and tactics refers to how forces actually fight. Are subordinate commanders trained to take the initiative or wait for orders? Do units engage or bypass enemy strong points? Do forces launch preparatory artillery barrages, or do they seek to maintain the element of surprise? The adoption of infiltration tactics, for example, jumpstarted the German offensive of March 1918 on the western front. Significantly, all three levels can vary independently of the current forces in being. Not only are there almost always several plausible tactical battlesystems and usage doctrines at any level of technology, but these battlesystems and doctrines generate a system of strategic interaction.

Warfare is fundamentally a “rock, paper, scissors” game. Choices are only dominant vis-à-vis other states’ choices. Historically, light missile infantry dominated heavy infantry, while heavy infantry armed with pikes or spears was invulnerable to heavy cavalry. But heavy shock cavalry always dominated light infantry, which lacked the ability to resist charges. If the opponent was fielding a heavy cavalry force, the best defensive countermeasure was a heavy infantry force. But if the opponent was armed with heavy infantry, the best countermeasure was light infantry. The optimal choice depends on the opponent’s decisions.

Combined arms warfare is the response to this fact, but the effectiveness of a specific balance of forces in a combined arms system is also subject to strategic interaction. Furthermore, the effectiveness of any weapons system and any combined arms system depends on how the forces are being used. Are the forces being used as raiders or as holders of territory? Are they being used offensively or defensively? The dynamics here are harder to illustrate, but consider this simple example: a strategic plan geared toward defeating an enemy army will work only if the opponent is willing to stand and fight. If the opponent chooses to use a Fabian strategy of avoiding conflict instead, the plan may come to naught.

**Territory, Conquest, and Victory: Unpacking the Assumptions**

This last point about the success of war plans raises a second problem with both the Van Evera and the Glaser and Kaufmann articles. Both articles assume that when seizing territory is easier, there is a greater propensity to use force. The logic is flawed. John Mearsheimer has argued convincingly that the expected rapidity of victory is the crucial determinant in decisions to use force. The problem is that seizing territory is

9. For an extended examination of warfare in the West that focuses on different strategies and interaction effects, see ibid., passim.
not synonymous with victory. Indeed, seizing and holding territory is neither necessary nor sufficient to win a war. Wars are won under two conditions. First, it is possible to win a war by effectively eradicating the ability of the other side to resist. With the exception of World War II in Europe, however, no war in modern history has ended as a result of the absolute destruction and occupation of a country’s territory. The second way to win a war is by either inflicting higher costs than the other side can accept or threatening credibly to do so. In many cases, this cost tolerance is not an objective measure, but rather a set of social constructions.

It is possible to inflict these costs without occupying territory. Indeed, for much of recorded history, the norm was for the losing side to concede the issue of the war following defeat in a major battle, even when this did not, in any significant and lasting way, undermine the losing side’s ability to wage war. In addition, historically, many wars have been won using raiding strategies in which control of territory is not sought. In ancient Greece, raiding was the dominant strategy. Alexander’s defeat of Persia was not the result of his ability to control territory. Even in the American Civil War, the devastating impact of William Tecumseh Sherman’s march to the sea had nothing to do with his ability to control territory.

To the extent that it is not necessary to control territory to win a war, offense-defense theory begins to break down. If defeating enemy armies and inflicting costs are major priorities, a fundamentally defense-dominant world, given a low enough force-to-space ratio, can lead to very rapid victories. In 1866 Prussia defeated Austria by winning a relatively indecisive victory at Königgrätz—the Austrians were able to withdraw in good order and link up with reinforcements from Italy—thereby causing a political crisis in Vienna. In 1870 the Prussians won a set of decisive victories against France by using turning maneuvers to wage an offensive strategy using the tactical defensive. These two very rapid and low-cost victories occurred in an era of extreme defense dominance in terms of tactical military factors. The explanation is that success in war and the ability to seize and hold territory are not coterminous.

The Implication of Complexity: The Problem of Post Hoc Justifications

Although one might argue that the offense-defense balance is worth examining on its own terms, offense-defense theory is often invoked as a concise way to expand the richness of systemic models of international relations, conflict, and even foreign policy. Thus much of the value of offense-defense theory derives from its contribution to building rich, powerful, and parsimonious explanatory and predictive theories.

In his article, Van Evera cites military factors, geography, social and political order, and diplomatic factors as causes of offense and defense dominance. Although he attempts to aggregate these into a single offense-defense measure, these variables are vague, too disparate to aggregate, and extremely dependent on subjective assessments. They produce a wide variety of conflicting theoretical predictions, most of which can be resolved only by empirical analysis.

It is difficult to avoid analytical bias in this process. For instance, Van Evera argues that “popularity of regimes probably aided offense before roughly 1800 and has aided defense since then. The reversal stems from the appearance of cheap, mass-produced weapons useful for guerrilla warfare—assault rifles and machine guns, light mortars, and mines. The weapons of early times (sword and shield, pike and harquebus, heavy slow-firing muskets, etc.) were poorly adapted for guerrilla resistance.” There were, however, many guerrilla campaigns before 1800. Longbows and crossbows were adequate guerrilla weapons. In this case, it is difficult to see how the coding can be done a priori. If Van Evera used the fact that guerrilla campaigns were more common and successful after 1800 as a basis for the judgment, then he may have conflated causes and outcomes. In any case, the argument is underspecified because the basic coding criteria are not explicit.

This lack of explicit criteria creates ambiguities in Van Evera’s article. Van Evera suggests that military factors favoring mass infantry enhance the offense, but he also stresses the limitations on offensive action imposed by the logistical demands of large forces. Cavalry forces, because they are expensive and hence limited in number, are argued to favor the defense despite their greater mobility, but tactically analogous armored forces (although with even larger logistical requirements than cavalry) are said to favor the offense. Mass infantry in the Napoleonic Era favors the offense. Mass infantry during World War I favors the defense. According to Van Evera, this distinction is the result of “lethal small arms, barbed wire, and trenches.” However, Borodino and Waterloo—Napoleon’s two major setbacks and the two most prominent battles where he fought a steady foe in a frontal assault—demonstrate the defensive power of mass infantry even in the early 1800s. Another example of this coding problem can be found by comparing the Van Evera and Glaser and Kaufmann articles. Whereas Glaser and Kaufmann argue that “the most widely agreed proposition is that improvements in mobility favor offense,” Van Evera argues that chariots, cavalry, and railroads—all

17. Ibid., p. 20.
21. Ibid., p. 17.
systems that enhance mobility—help the defense. The other variables cited by Van Evera and Glaser and Kaufmann—the impact of geography, social and political order, and diplomatic factors—are even less susceptible to clear coding.

A larger problem is Van Evera’s belief that these variables can be aggregated. This is troubling from a methodological perspective. Van Evera does not present any conceptual explanation for how he actually measures the offense-defense balance in each area. Instead, he presents a laundry list of things to look for. Nor does he explain the relative weights he uses in aggregating his offense-defense variables. This leads to such confusing passages as, “Sometimes technology overrode doctrine, as in 1914–18 and in 1945–91 (when the superpowers’ militaries embraced offensive doctrines but could not find offensive counters to the nuclear revolution). Sometimes doctrine shaped technology, as in 1939–45, when blitzkrieg doctrine fashioned armor technology into an offensive instrument.” Without a set of contingent generalizations about the conditions that define the “sometimes,” the theory assumes what it ought to demonstrate.

Offense-defense theory represents what Giovanni Sartori called “concept misformation.” By defining the balance as being a function of a vast, unrelated grab bag of conditions and variables, Van Evera and Glaser and Kaufmann have created a situation where the empirical referents become merely a menu of items to choose from to justify a preexisting assessment of what the offense-defense balance is at a given point in time. This problem is exacerbated because most observers begin their research with significant knowledge and preconceptions about what the offense-defense balance was during the particular period they study. Familiarity with the cases almost certainly leads to bias in interpretations. Glaser and Kaufmann compound this problem with their methodological approach. They assume that if you throw the right variables together and do some net assessment, the offense-defense balance will emerge. This sort of naked empiricism does not advance the cause of theory building.

A Parsimonious Addition?

It might be possible for a historically knowledgeable and methodologically sophisticated scholar to develop a comprehensive model of the causes of the offense-defense balance. The extraordinarily complex resulting model would not, however, be parsimonious.

A fully specified model of the sources of offense and defense dominance would take into account the interaction of different possible battlesystems and different usage options. This process would create a broad typological theory. Then by examining the cost of the competing system, we might be able to derive a crude offense-defense

24. Ibid., p. 18.
26. Ibid., p. 1039.
balance for a given typological space. This result would then need to be validated by some sort of empirical analysis across cases. Given that the process of operationalizing the balance is unwieldy, involving complicated theorizing and budgetary and campaign analysis, it is reasonable to question whether the offense-defense balance adds sufficient richness and explanatory power to justify the very significant loss in parsimony.

Conclusion

Offense-defense theory is methodologically flawed and conceptually muddled. Although Van Evera and Glaser and Kaufmann push offense-defense theory forward, it is time to ask whether offense-defense theory in fact moves the field forward, or whether it represents instead the security studies' version of the emperor's new clothes.

Defenders of offense-defense theory will likely make several responses. First, they will correctly argue that I fail to address the perceptual variant of offense-defense theory. However, if the offense-defense balance is not an objective or structural condition, but instead resides purely in the realm of perceptions, then it ought to be integrated into a cognitive-processes framework rather than held apart as a special sort of (mis)perception.

Second, defenders of offense-defense theory will claim that I overstate the difficulty of operationalizing the balance or that I overemphasize the complex interaction effects of tactical battlesystems, doctrines, and usage decisions. If that is the case, however, then I would simply ask them to demonstrate the operationalization process in a systematic, reproducible manner. Until that time, I will remain a skeptic.

Third, offense-defense theorists will argue that I overstate the problem with relying on seizing territory as a measure of offensive success. They may claim that substituting “victory” for “conquest” or “seizing territory” is a simple change that does not harm the theoretical construct. I would argue, however, that once one breaks the link between ease of seizing territory and victory, one is left with the fact that none of the empirical indicators operates in a consistent fashion. This is not a semantic distinction. Rather, all the “causes” of the offense-defense balance are derived from the ease of seizing territory. Once that link is broken, the entire logic of the argument is questionable.

Finally, and linked to the previous point, offense-defense theorists will claim that even if my arguments about interaction and mobile warfare are valid, the theory is still useful because it explains the conditions under which breakthroughs occur. These breakthroughs, offense-defense theorists might argue, are a prerequisite for any kind of military victory, and they require the sort of frontal, attrition assaults the relative costs of which offense-defense theory claims to measure. The problem with this argument is threefold. First, although the costs of breakthrough may be relatively higher or lower, breaking through on a narrow front may not raise the costs significantly for the campaign as a whole. Second, this sort of breakthrough implies a situation of high force-to-space ratios and a continuous front. Empirically, this is not a common condition. Third, breakthroughs do not necessarily require attrition and frontal assaults. Fronts can be broken by infiltration tactics and can usually be outflanked by a sufficiently imaginative foe.
So what can be done? Can offense-defense theory be saved? The short answer is, no. Offense-defense theory has too many critical and fatal flaws. If we cannot “uninvent” offense-defense theory, then we must be very cautious about how we use it. There is an unfortunate tendency in the field to believe that offense-defense theory is a cheap and easy way to add predictive power to an explanatory model. In reality, the issues raised by offense-defense theory are extremely complex and difficult to parse effectively. The theory creates more conceptual holes than it fills, and should come with a strong warning label attached.

—Bernard I. Finel
Washington, D.C.

To the Editors (Stacie E. Goddard writes):

Stephen Van Evera’s and Charles Glaser and Chaim Kaufmann’s recent contributions to International Security are welcome expansions of offense-defense theory. Both articles recognize that although hypotheses presented in this literature are intuitively and empirically plausible, offense-defense theory has suffered from a lack of methodological rigor: definitions are often tautological; the variables of offense and defense dominance are continuously conflated with other factors significant to international relations theory; and at times hypotheses seem to be nothing more than “folk theorems” derived from the popular case of World War I. Unfortunately, while both articles attempt to address and overcome these critiques, neither satisfactorily resolves the methodological problems mentioned above. Most notably, these authors do not distinguish the offense-defense balance from factors such as the balance of power and military skill. This in turn leaves them vulnerable to tautological propositions, overdetermination, and difficulties with empirical measurement and testing.


I begin by examining Van Evera’s article, arguing that his conceptualization of offense dominance as the “ease of conquest” confuses the offense-defense balance with the probability of success in war. This conceptual problem has serious ramifications for the explanatory hypotheses and empirical evidence he brings to bear in his theory. I then turn to Glaser and Kaufmann, noting that while the authors are cognizant of previous methodological critiques, the theorists cannot avoid conflation with other variables in their framework. Specifically, their broad definition of the offense-defense balance, combined with their assumption of “optimal doctrine,” is insufficient to distinguish offense dominance from either power or doctrine. Following these critiques, I conclude by offering definitions and suggestions for empirical testing that might help avoid these methodological problems in future research.

The Ease of Conquest and the Probability of War

In “Offense, Defense, and the Causes of War,” Van Evera argues that “war is far more likely when conquest is easy, and that shifts in the offense-defense balance have a large effect on the risk of war.” He offers ten explanatory hypotheses, including the temptation to strike first and increased incentives for expansionism, that further link his definition with outbreaks of war. After testing these explanatory propositions against three periods in history (Europe since 1789, ancient China during the Spring and Autumn and Warring States periods, and the United States since 1789), Van Evera concludes that “offense-defense theory has the attributes of a good theory,” explaining large amounts of international history with a single variable.

Although many of the theoretical and empirical propositions are compelling, serious methodological flaws detract from Van Evera’s argument. First, his definition of offense dominance—“conquest is easy”—confuses offense dominance with a host of other variables, most notably with the balance of power. Simply put, Van Evera has defined offense dominance in terms of war outcomes, confusing the offense-defense balance with the probability that an attacking state will prevail in the event of war. To say that an attacking state has a high probability of defeating its opponent says nothing, however, about the relative efficacy of offensive and defensive operations per se. One can easily imagine a scenario in which defensive strategies have an advantage, yet the attacking state prevails because of superior logistical support, deeper economic resources, or an overwhelming advantage in the number of forces. For example, conquest certainly appeared “easy” for the North at the close of the American Civil War. This was not a function of the strength of offensive versus defensive strategies, but of the greater social, economic, and logistical support of the population. In short, in order to

4. Ibid., p. 41.
5. Also making this critique are Glaser and Kaufmann, “What Is the Offense-Defense Balance?” p. 70; however, I argue that they too are vulnerable to these criticisms.
distinguish the offense-defense balance from the balance of power, Van Evera needs to cast his definition in relative terms (i.e., the value of attacking compared to the value of defending), rather than focus on the absolute value of attacking for a state.

These conceptual problems are compounded in Van Evera’s explanatory hypotheses. For instance, he asserts that “when conquest is easy, aggression is more alluring: it costs less to attempt and succeeds more often.”7 Therefore “resources are more cumulative when conquest is easy. . . . As a result, gains are more additive.”8 It may be the case that conquests are self-reinforcing; however, this hypothesis cannot logically be derived from an assessment of offensive and defensive strategies. Although the offense-defense balance can tell us the relative costs of attacking versus defending, it measures neither the absolute probability of success nor the absolute value of conquest.

Furthermore, Van Evera’s causal explanations come close to tautologies: it often seems he is arguing that when conquest is easy or perceived to be easy, states will attempt to conquer. Indeed, myriad variables—including state behavior—are subsumed in his conceptualization. According to Van Evera, “Military technology and doctrine, geography, national social structure, and diplomatic arrangements (specifically, defensive alliances and balancing behavior by offshore powers) all matter” in determining the offense-defense balance.9 One is left wondering what factors could be excluded from this definition to show the causal autonomy of the offense-defense balance. Moreover, after subsuming all of these behavioral variables into his conceptualization, he uses the offense-defense balance to explain behavior among states. For instance, after the Crimean War “the power of defenders fell dramatically because defense-enhancing diplomacy largely broke down.”10 In the end, Van Evera is using the behavior of states (eschewing defensive alliances in favor of offensive diplomacy) to explain the behavior of states (offensive diplomacy until 1871).11

Finally, these methodological flaws are strikingly evident in his empirical accounts. Most important, Van Evera makes no attempt to avoid conflating the offense-defense balance with the balance of power or military forces. For example, he argues that “during 1792–1815 the offense was fairly strong militarily, as a result of France’s adoption of the mass army (enabled by the popularity of the French government).”12 Although it is certainly plausible that Napoleonic doctrine and tactics related to mass armies favored the offense, the size of the army itself is indeterminate of the offense-defense balance. Arguably, Napoleon’s mass army would have been more effective defending French soil, rather than searching out offensive campaigns across Europe.13

8. Ibid., p. 8.
9. Ibid., p. 6.
10. Ibid., p. 28.
11. Van Evera avoided this tautology in “The Cult of the Offensive and the Origins of the First World War,” International Security, Vol. 9, No. 1 (Summer 1984), pp. 58–107. In this article, alliances are a clearly distinguished dependent variable. The independent variable is perception of offense dominance, and thus Van Evera does not incorporate states’ behavior into his explanation.
That France could deploy a mass army tells us a lot about the power of the centralized state, but very little about the offense-defense balance.

In sum, while Van Evera argues that offense-defense theory offers both unique and plausible predictions, methodological flaws with his approach undermine this claim. By subsuming power, military skill, diplomacy, doctrine, social structure, and domestic political structure into his conceptualization of the offense-defense balance, Van Evera cannot argue that the offense-defense balance has more explanatory power than any of these variables taken separately. Needless to say, a much narrower definition and systematic hypotheses are necessary before these claims to progress can be empirically tested.

Power, Skill, and Strategy: The Offense-Defense Balance and the Optimality Assumption

In “What Is the Offense-Defense Balance and Can We Measure It?” Glaser and Kaufmann are particularly concerned with previous methodological critiques of offense-defense theory. They note that critics have questioned the utility of the theory on the grounds that “the foundations of the theory are underdeveloped” and that “the theory contains inherent flaws, the most serious of which is that the offense-defense balance cannot be measured because the outcomes of war are so uncertain.”

In response to these criticisms, Glaser and Kaufmann state that the offense-defense balance should be defined relatively: “the ratio of the cost of the forces that the attacker requires to take territory to the cost of the defender’s forces.” They argue that this definition, combined with six key assumptions, allows offense-defense theorists to avoid conflation with other variables. Moreover, the theorists contend that the offense-defense balance can be measured by using the analytical toolbox of military net assessment, thus allowing empirical tests of the theory.

Clearly, distinguishing the offense-defense balance from other factors is a crucial task for Glaser and Kaufmann. They state that they “envision offense-defense theory as a partial theory of military capabilities. . . . A more complete theory would include two additional variables: (1) power, measured in terms of relative resources; and (2) what we term ‘military skill,’ that is, a country’s ability to effectively employ military technology.”

Indeed, Glaser and Kaufmann do not claim that the offense-defense balance is the only or even primary determinant of military outcomes, but rather “each of these three variables has the potential to overwhelm the others in certain circumstances.”

Although Glaser and Kaufmann realize that distinguishing the offense-defense balance from power and skill is important, they fail to do so adequately in their methodological framework. First, their adoption of a broad definition of the balance subsumes competing factors, such as force size and even nationalism. Although these are critical factors in determining the outcome of a war, they are more likely reflective of the power

15. Ibid., p. 46.
16. Ibid., pp. 48-49.
17. Ibid., p. 49.
of a centralized state, and the ability of the state to effectively mobilize resources and forces, than they are of the relative efficacy of offensive and defensive strategies. As I argue above, it is plausible that an attacking state with popular support and a mass army will defeat a country lacking these attributes. This, however, does not tell us the relative value of offensive and defensive strategies for either of the actors involved.

Glaser and Kaufmann face similar methodological problems when trying to distinguish the offense-defense balance from military skill. They argue that “the offense-defense balance should be assessed assuming optimality—that is, countries choose the best possible strategies and force postures for attack and defense. Offense-defense theory requires this assumption because it focuses on the effects of the constraints and opportunities presented by the international environment.”18 The problem with this assumption is that it invariably leads to a post hoc, ergo propter hoc fallacy—we can only ascertain the optimal strategy after observing which strategies succeeded during a war. By measuring the balance in terms of successful strategies, we can neither determine the balance ex ante nor satisfactorily separate the balance from power and skill.

For instance, consider the offense-defense balance during World War I. Most would concur that the balance heavily favored the defense, both before and during the war, although this balance was misperceived by statesmen and military leaders alike.19 Militarily, this defensive advantage is epitomized by the 1916 Battle on the Somme, an attack launched by the British on July 1, 1916, and lasting through November of that same year. Over this five-month span, and at a cost of approximately 500,000 casualties (the largest number ever of British casualties in battle), the British were able to move the front only seven miles. Two years later, however, on March 21, 1918, the Germans achieved a massive breakthrough on the same terrain, using the same type of weapons available to the British in 1916. After breaking through the juncture of French and British troops, the Germans employed a creeping barrage and infiltration tactics to gain more ground on the first day of attack (approximately 140 square miles) than the British had in 140 days. The German strategy in World War I would therefore be considered by Glaser and Kaufmann to be the optimal doctrine. In fact, using these criteria would mean that World War I was “objectively” offense dominant.20

18. Ibid., p. 46.
Although this conclusion may seem bizarre, I have trouble seeing how Glaser and Kaufmann would refute it. A possible response might be that while my example focuses on tactics, they claim to measure the offense-defense balance at the strategic level of conflict. This argument is problematic for two reasons, however. First, as Glaser and Kaufmann note, “a change that shifts the balance in a given direction at one level will usually also shift it in the same direction at all higher levels.” Therefore “a change that makes tactical offense harder will usually also make operational offense harder, which in turn makes strategic offense more difficult.” Indeed, tactical innovations at the Somme in 1918 made offense easier at the operational and strategic levels for the Germans, and later the Allies. Furthermore, even if distinguishing tactical from strategic shifts in the offense-defense balance is theoretically possible, Glaser and Kaufmann provide insufficient guidance on how this could be accomplished. In fact, their own net assessment techniques rely on both tactical and strategic calculations to measure the offense-defense balance.

In sum, Glaser and Kaufmann’s assumption of “optimality” neither controls for military skill nor allows for ex ante assessment of the offense-defense balance. Indeed, the optimality assumption means that one would inherently code periods in terms of skill, measuring the offense-defense balance in terms of the most successful strategy employed. In further research, this dilemma could be addressed by assuming symmetrical doctrine. Obviously, whether doctrine is sufficiently symmetrical is ultimately an empirical question. This assumption, however, would allow theorists to compare the efficacy of offensive and defensive strategies while controlling for skill and doctrine.

**Conclusion**

Although offense-defense theory is riddled with methodological problems, the hypotheses put forth by these theorists are both empirically plausible and policy relevant. The Van Evera and Glaser and Kaufmann articles deserve attention and scrutiny for these reasons. Critical revision of the theory is clearly in order; to this end, I offer two suggestions.

First, a definition of the offense-defense balance that avoids subsuming power and doctrine needs to be constructed. The balance cannot be represented as the probability of taking territory, and should not incorporate the absolute value of an attack. An example of such a definition is one that describes offense dominance as a situation in which it costs less in terms of lives and territory to attack a state than it does to defend against it. Analogously, defense dominance would imply that attacking costs more in terms of lives and territory than defending against an attack. The offense-defense balance is represented as the cost difference between the two. Not only is this particular definition a comparison of the relative efficacy of the strategies, it says nothing about the ultimate outcome of a war, and thus avoids incorporating power into the concept.

22. Ibid.
Moreover, this conceptualization allows that the offense-defense balance might vary from state to state, given that it is framed in terms of a subjective utility function.

Second, empirical tests of the offense-defense balance must be far more rigorous. Coding periods as offense or defense dominant by looking at the outcomes of war does not tell us about the efficacy of strategies and could easily be representative of balance-of-power factors. To avoid this, one might consider cases where states faced the possibility of conflict with each other over an extended period of time. Finding time periods in which the offense-defense balance varied, while power and doctrine was relatively constant, would be difficult; however, this would allow the variables of power and doctrine to be controlled.

—Stacie E. Goddard
New York, N.Y.

Stephen Van Evera Replies:

James Davis, Bernard Finel, and Stacie Goddard raise a number of questions about offense-defense theory. Here I focus on three that seem most important, briefly address five others, and let the rest pass without comment, except to express a general dissent from their arguments.

First, can we characterize specific military technologies or force postures as defensive or offensive? Or is everything dependent on the context of combat?

Bernard Finel takes the latter view, arguing that “the offense-defense balance is . . . an outgrowth of strategic interaction.” He much overstates a good point. As he notes, the capacity of a military force can depend on the forces and strategies of its opponent. Some forces are better at attack than defense against some opponents, while being better at defense than attack against others. But many forces are inherently optimized for offense or defense, in a way that applies across opponents.

A secure nuclear deterrent is fundamentally defensive. It makes its owner essentially unconquerable. At the same time, it cannot conquer other states that possess a secure deterrent. These facts apply regardless of the opponent’s strategy. The defensive character of the nuclear revolution, which stems from these realities, is the defining feature of modern international relations. It is the single most important aspect of post–World War II and post–Cold War international affairs.

Because nuclear deterrents are defensive, forces designed to counter them are essentially offensive. These forces include offensive strategic nuclear counterforce systems (e.g., accurate intercontinental ballistic missiles and strategic antisubmarine warfare systems) and area ballistic missile defenses.

Modern guerrilla war has defended many countries and conquered none. It is a fundamentally defensive form of warfare. States would pose little threat to one another if all relied on citizen guerrilla defenses.

The accurate repeating rifles, machine guns, barbed wire, railroads, and entrenchments of the western front in World War I were fundamentally defensive, as the

repeated failed offensives of 1914–17 demonstrated. They were eventually overcome: first by the German infiltration tactics in 1918, and later by German blitzkrieg concepts of armored war. But for three years they gave dominance to the defense.

Many other developments in military history can be characterized the same way. The weapon and tactics of the late Middle Ages in Europe advantaged the defense; the gunpowder revolution then overcame these weapons and tactics, restoring the offense; innovations in fortification by Vauban and others restored the defense in the late seventeenth century; and Napoleonic warfare by popular mass army then strengthened the offense. In the twentieth century, German armored blitzkrieg tactics also bolstered the offense. The effects of these modes of warfare varied only modestly with the nature of their opposition. They had an inherent propensity to ease the defense or the offense.

In sum, technology and force posture do, on important occasions, have innate defensive or offensive properties and implications. Nuclear weapons are the most important recent example, but they are only one among many.

Second, does offense-defense theory lack parsimony? Does it commit the sin of explaining by complexitying?

Finel claims that it does. Its independent variable—the ease of conquest—includes factors drawn from the military, diplomatic, geographic, and social spheres. These factors are, says Finel, a “vast, unrelated grab bag of conditions and variables.” The use of such an unwieldy variable leads to a “very significant loss in parsimony.”

Finel has things backward. Offense-defense theory is elegant. It is parsimonious. It orders and thereby simplifies a previously disordered melange of phenomena.

To judge Finel’s charge, we first must ask: What is parsimony? What provides it, and what detracts from it?

A theory is not shown to lack parsimony simply by demonstrating that its concepts include a diverse range of lesser-included concepts, because this is true of every concept. All concepts are aggregations of lesser concepts. For example, national power is a concept that aggregates national military power, economic power, and the power to marshal allies. But thinking about strategy would be more complicated, not simpler, if we dropped national power from our lexicon and discussed only its component parts. Military power, a main component of national power, is an aggregate of airpower, naval power, and ground power, as well as of material resources, skill, and willpower. Discussion would be far harder if we had to address these components separately each time military questions were at issue. Airpower, a component of military power, aggregates the quality and quantity of aircraft, the quality and quantity of air force personnel, and the quality of air doctrine. The quality of aircraft in turn is an aggrega-

---

tion of the speed, range, payload, maneuverability, stealthy characteristics, avionics, and durability of the aircraft. At every level, we find that concepts are composed of more concepts from the level below.

My rule of thumb is to judge new concepts by asking if they are theoretically useful. In political science this standard usually requires that concepts somehow correspond to phenomena in the real world. They must fit the way things work, or the way we think. A concept that fails to do this is artificial and only clutters discussion. But a concept that combines other concepts while also capturing reality simplifies discussion.

When considering national security problems, leaders often ask if others can conquer them or if they can conquer others. These questions lie at the core of many past and present foreign and security policy debates. And in asking if conquest is possible, leaders aggregate the same military, diplomatic, geographic, and social factors that I aggregate to capture the ease of conquest. Leaders aggregate because they must; because these factors together decide if they can conquer or be conquered. Aggregation is not easy and cannot be precise, but leaders do it because otherwise they cannot understand their national security situation. As Bismarck planned his wars of German unification, he recurrently asked if Germany was in a position to wage aggressive war successfully. To answer, he had to weigh military, diplomatic, and geographic considerations in some combination. Later, Germans who warned of German insecurity weighed these same factors together. Americans did the same when assessing their national security during the Cold War. Former President Herbert Hoover opposed the U.S. troop deployment to Europe in 1951 because he thought conquest was difficult: “This Hemisphere can be defended from Communist armies come what will. . . . Communist armies can no more get to Washington than any allied armies can get to Moscow.”4 Hoover had to aggregate military, geographic, and diplomatic factors to reach this conclusion. Oppositely, the authors of NSC-68 feared in 1950 that the United States faced a grave threat to its security, in large part because they believed that conquest was easy. They reached this conclusion by aggregating both military and diplomatic factors.5

The concept of ease of conquest, then, corresponds to the way policymakers think. It captures the way they organize the world. If it did not, it would be a complicating distraction. But it does. And in so doing, it simplifies our discussion of security problems and policies.

Offense-defense theory achieves other simplicities as well. It argues that a number of important war causes—expansionism, fierce resistance to others’ expansion, first-strike advantages, windows of opportunity and vulnerability, fails accomplis, negotiation failures, secrecy, arms races, and “chain gaining” in alliances—that were once

5. See NSC-68, excerpted in Thomas H. Etzold and John Lewis Gaddis, eds., Containment: Documents on American Policy and Strategy, 1945–1950 (New York: Columbia University Press, 1978), pp. 414, 416 (suggesting that nuclear weapons are offensive), and pp. 427, 430 (suggesting that states tend to bandwagon with threats). Also relevant is p. 396, suggesting that the Soviet empire was vulnerable to Western offensive action, for essentially social reasons.
viewed as independent stem from a single cause. This simplifies the problem of power and war. We see that a number of disparate dangers are fed by a single taproot.

In short, offense-defense theory explains a wide range of phenomena with a parsimonious framework. Far from complexifying, it streamlines our understanding of the war problem.

Third, is offense-defense theory testable? James Davis complains that my formulation of offense-defense theory is too imprecise and incoherent to be tested. Specifically, he notes that I consider two variants of offense-defense theory—an objective variant and a perceptual variant—but I infer and test only one set of predictions from these two variants. Instead, he argues, each variant must have its own distinct set of predictions. Davis makes a good point, but his conclusion is overdrawn. He shows that I failed to explain myself, but does not show that offense-defense theory is untestable or is otherwise flawed.

Let me clarify the confusion that Davis rightly identifies. Offense-defense theory has an objective and a perceptual variant. The objective variant frames the effects of the actual offense-defense balance. The perceptual variant frames the effects of the perceived offense-defense balance. Both the objective and the perceived offense-defense balances are indicators of the other: the objective offense-defense balance influences—and thus indicates—the perceived offense-defense balance; and the perceived offense-defense balance is influenced by—and thus indicates—the objective offense-defense balance. Therefore both the objective and the perceptual variants of offense-defense theory make predictions about the correlates of both the objective and the perceptual offense-defense balances. I conflated these two sets of predictions in my article, offering a single unified forecast about how things would appear if the offense were dominant or if it were believed dominant. I did this because it works in this instance: both variants of the theory make parallel predictions about both objective and perceived reality. We need not distinguish these forecasts because they are essentially the same. Two variants of a theory usually produce two divergent sets of predictions, but not in this case.

What do we conclude when the objective and perceived offense-defense balances differ, as in 1914, when the objective balance favored the defense and the perceived balance favored the offense? Both variants of the theory make stronger predictions about the correlates of the perceived balance than the objective balance, so tests that look to the perceived balance are stronger. For example, in the 1914 case both variants predict that perceptions of offense dominance should correlate with war, even if objective realities favor the defense. The perceptual variant predicts this simply because it deals only with the perceptions, and puts the objective balance aside. The objective variant predicts this because the impact of the objective balance is translated into outcomes through its effect on the perceived balance, as follows: objective offense-defense balance --> perceived offense-defense balance --> outcomes. The hypothesis on the right (perceived offense-defense balance --> outcomes) should operate even when the hypothesis on the left (objective offense-defense balance --> perceived

offense-defense balance) does not. Hence even the objective variant forecasts a more certain correlation between perceptions of offense dominance and war than between objective offense dominance and war.

I saved my readers this detail because I feared that their eyes might glaze over. But I should have explained it, and Davis is right to complain that my failure to explain is confusing. He is wrong to claim, however, that offense-defense theory is somehow flawed or untestable. I failed to explicate my logic, but I think that logic is sound.

I close with remarks on five other criticisms made by Davis, Finel, and Goddard.

Finel argues that seizing territory does not confer victory in war.7 He seems to think I believe otherwise, but I agree with him. My view is that the seizing of territory requires victory (not the other way around). Clearly, without a victory of some kind it is impossible to seize and hold another state’s territory.

Finel quarrels with my coding of the history of military technology and strategy. He doubts that guerrilla war grew easier with the development of mass-produced modern small arms after 1800, and that accurate repeating rifles, machine guns, and barbed wire made frontal assaults more difficult in the late nineteenth and early twentieth centuries.8 In so arguing, he takes on many historians in addition to myself. I concur, however, that such questions are not open-and-shut, and we need a detailed study of the history of the offense-defense balance in warfare to help resolve such disputes.

Stacie Goddard argues that I failed to distinguish the offense-defense balance from the balance of power.9 I certainly meant to distinguish them, and believe I did. In footnote 1, I suggest that the offense-defense balance could be measured by looking at the probability that a determined aggressor could conquer a target state with comparable resources. In other words, this measure asks how often conquest occurs where the balance of power cannot account for the outcome, because the winner starts with no marked resource advantage. It should be clear from this that I am not running together the offense-defense balance and the balance of power.

Goddard and Davis contend that I use the behavior of states to explain the behavior of states.10 I plead guilty. In fact, the behavior of states often explains the behavior of states. Europe’s continental powers reach for hegemony when Europe’s offshore balancers (Britain and the United States) are in an isolationist mood, and are more cautious when the balancers are active. If that’s how the world works, shouldn’t we say so? I don’t see a problem.

Davis objects that I define terms in ways that will confine others’ analyses.11 But of course others are free to adopt the definition that best helps them answer their questions, just as I did. I cannot stop them and would not want to.

8. Ibid., pp. 184–185.
Davis, Finel, and Goddard have raised important issues that deserve further attention, but offense-defense theory stands up to their criticisms.

—Stephen Van Evera
Cambridge, Massachusetts

Charles L. Glaser and Chaim Kaufmann Reply:

We appreciate the opportunity to respond to the issues raised by Bernard Finel and Stacie Goddard. We believe that there is little real disagreement between us and Goddard—she makes a number of sound arguments, but on some points apparent differences between her position and ours result from her misunderstanding of our views. In contrast, we think that Finel's wholesale condemnation of the entire offense-defense research program is at least premature, based on failures to understand both the purposes of offense-defense theory and the requirements for testing it. Goddard's and Finel's most important points relate to the status of offense-defense theory as a structural theory of international behavior. Therefore we address these issues first, and then turn to several separate issues raised by one or the other correspondent.

As a structural theory, offense-defense theory attempts to predict states' behavior by focusing on the constraints and opportunities presented by their environment. Important constraints include the offense-defense balance and the distribution of resources (power), while among the choices that the theory seeks to predict are decisions about military doctrine and force posture, as well as whether to form alliances and fight wars.

Structural theories of international politics can incorporate various kinds of constraints, some of which are stricter than others. System structure, as defined by Kenneth Waltz, excludes properties of units (typically states). In Waltz's narrow definition, structure consists only of properties that emerge from the relationships of the units to one another and that no individual state can change—international anarchy, for example. Constraints that emerge from purely material facts, which can be properties of states but which states cannot change or evade, we can call material structure; with limited exceptions, geography and weather would qualify. Theories based on these hard constraints are, however, often not satisfying, because they miss too much of interest in most areas of international behavior.

Most structural theories therefore employ a “softer” definition of structure that includes any constraints that states cannot change or evade within the time scales they are likely to consider in planning foreign and defense policy. These may include social facts such as a state's form of government or its level of scientific achievement at a given time. Measurements of power, for example, must include not only a state's purely material resources but also the capacity of the state apparatus to extract resources from

2. Appropriate time scales for different types of foreign policy decisions may vary. For offense-defense theory, which is principally concerned with understanding decisions in peacetime about planning for deterrence or for fighting future wars, we think a relevant time scale is often from several years to ten years, although there could be variations—for example, if a state could be highly confident that it would face no threats for a longer period.
society for military use. Our measure of the offense-defense balance includes whether the state—and its adversaries—are nation-states or multinational empires, as well as many aspects of technology. Including certain unit-level factors does not erase the difference between structural and nonstructural theories. What all structural theories exclude are explanations of the actual decisionmaking process, including the possibility of flaws of perception and judgment, and information about state preferences that is based on their unit-level characteristics.

To understand the impact of constraints, structural theories, including offense-defense theory, must assume that states’ policy choices are broadly optimal or rational; subject to the constraints they face, states make effective policy choices for maximizing their interests. Theories that do not assume optimality must include a theory of suboptimal state decisionmaking. In such theories much of the explanation of state behavior is often attributed to domestic political competition and/or to flaws in the decisionmaking process; the impact of environmental constraints on policy choices is weaker and harder to isolate (although those same constraints may still exert a powerful influence on the eventual outcomes of policy choices).

The optimality/rationality assumption in structural theories is useful in three ways: it can provide guidance for making policy; it establishes a baseline against which states’ policies can be compared to determine whether they are flawed; and it can help assess the likely impact of flawed policies. By providing a baseline, the optimality assumption in offense-defense theory enables us to separate military skill from the balance. The balance is measured assuming that all countries have high levels of military skill. Choices that diverge from the baseline are suboptimal and indicate low skill.

Goddard objects to our inclusion of an optimality assumption in offense-defense theory, arguing that it leaves us unable to separate military skill from the offense-defense balance. The problem, as she sees it, is that the optimal strategy can be determined only “after observing which strategies succeed during a war,” which means that we cannot determine the balance ex ante. We agree that the balance should be measured ex ante—this is required by the theory. Goddard’s objection is based on an exaggeration of the standard of optimality required for purposes of the theory. She apparently takes optimality to mean the absolute best choice the state could make given not just the (limited) information available at the time but also the information that would be provided by future actions. With this understanding of optimality, it would

3. Constraints could include facts that decisionmakers theoretically could change, but only at costs they would likely consider unacceptable. For instance, a multinational empire could convert itself into a nation-state by giving up its imperial territories and subjects, but this would likely violate the rulers’ sense of the identity of the state, as well as reduce its resources.
4. Except in formal rational choice-oriented work, the assumption is not usually that actual state decisionmaking processes meet normative standards of rationality, but rather that decisionmakers act “sensibly” given the information available to them; decisionmakers usually make choices that do not vary much in substance from those that would be made by purely rational actors.
indeed follow that optimal force posture, doctrine, and strategy cannot be known until they are revealed by war outcomes, and that ex ante measurement of the offense-defense balance would be impossible. However, this is not what we mean by optimality; the standard that we employ is that states choose optimally “within reasonable limits of analysis” given the information available to them at the time; to impose a stricter standard would be unreasonable for a theory intended to predict actual behavior. Our standard is the same optimality/rationality assumption that appears in most structural theories.

Based on our understanding of structure, we also reject two of Finel’s key points. First, he argues that “the offense-defense balance is not a structural variable. Rather, it can be influenced by immediate decisions about deployments and employment strategies.” This is incorrect. The offense-defense balance is a constraint, not a measure of the effectiveness of actual deployed forces for either offense or defense. It answers the question: How secure can states be, assuming that both they and their opponents make optimal choices? The offense-defense balance, in combination with power, determines how well a state can do; state decisions in combination with structural constraints determine how well a state will actually do. Suboptimal decisions reduce the state’s military capability compared to the best that it could be, but do not influence the balance itself. For example, if one state deploys nuclear weapons in vulnerable basing modes, then an attacker’s prospects for significantly limiting damage (and therefore for a successful offensive attack) will be much greater than if the state had made better deployment decisions. However, the offense-defense balance for this example remains defined by the best retaliatory capability that the state could achieve given both sides’ resources and available technology.

Second, Finel argues that the offense-defense balance is not a structural variable because it is an “outgrowth of strategic interaction,” so that “the optimal choice depends on the opponent’s decisions.” In response, we would first like to point out that a great many military policy decisions, especially at the levels of doctrine and force posture, are pure optimization problems involving no interaction. Before World War I, all armies would have been better off deploying more machine guns and less cavalry—regardless of what anyone else did. Between the two world wars, all navies would have been better off investing less in battleships. Finel’s observation that ancient armies used widely varying force combinations does not imply, as he suggests, that the best force

7. Actually, optimal strategy in Goddard’s sense cannot be reliably determined even after a war. Ex post we know that the victor’s choices were successful, but not necessarily whether they were optimal. The victor could have won despite suboptimal strategy because of even worse suboptimal choices made by the loser. Alternatively, even if the loser’s choices were optimal and the victory was the result of superior power, it is possible that the victor could have succeeded even more easily with a better strategy that was not tried.
9. This does mean that when information that would affect states’ strategies is unavailable, decisionmakers’ estimates of the offense-defense balance may differ from the estimates they would have made had they had the additional information. A possible example is the difficulty in 1939 of estimating whether atomic weapons would be developed in time for use in World War II.
posture generally depends on the force posture of the opponent. Nearly all of the most successful ancient armies were based on cores of heavy infantry, with lesser investment in cavalry and light troops. When armies were organized on other principles, this was usually because of limitations imposed by social systems or by terrain, not by the nature of the opponent's army. When structural constraints are strong, not only doctrinal but also some wartime strategy choices can become noninteractive. For example, at the start of World War I, all European states would have been better off if they had scaled down their offensive plans, regardless of what other states did.

Strategic interactions do occur. When they do, whether Finel's point is valid depends on exactly what we mean by "interaction." One sense of interaction would mean that each state must optimize its military doctrines and strategies not only subject to those structural constraints that influence it directly, but also subject to the knowledge that their opponents will also be attempting to optimize their own choices subject to the constraints facing them. The offense-defense balance is estimated assuming that both sides do the best they can, each knowing that the other side is also doing the best it can. For example, a state attempting to develop an effective nuclear damage-limitation capability against a particular opponent would have to take into account not only the technical, geographical, and other limits on its counterforce capabilities, but also its opponent's best options for improving the survivability of its strategic forces. If missile accuracy is low, the opponent can deploy survivable forces with little effort, and the balance will strongly favor defense. If accuracy is high and the opponent lacks the technology for survivable launchers such as mobile missiles and nuclear-powered submarines, it will have to spend much more to achieve a robust retaliatory capability, and the balance will be more favorable to offense. In any case, the resulting measure of the balance is fully structurally determined (provided that states do not make suboptimal choices). To the extent that each state responds not only directly to structural factors but also to the other's behavior, it is simply incorporating additional structural effects mediated indirectly through their pressure on the opponent.

A second sense of "interaction" would be a situation in which structural constraints exert only a weak influence on strategy choice, so that each side has two or more options that are equally good from a structural point of view, and that are different enough from each other that the best counterstrategies against each are quite different. Thus success for each side would depend in large part on correctly guessing the other side's choice, but there is no way to guess except by having some insight into the opponent's decisionmaking process, or by luck. In Finel's terms, this would be a true "rock-scissors-paper" situation where there is no one optimal choice, which would indeed mean that the offense-defense balance could not be measured exactly, but would spread out into a band of uncertainty whose width would reflect the impact on war outcomes of different combinations of a state's right or wrong guesses about its opponent's choices.

Such situations do occur in war, although they are more common at the operational and tactical levels, where individual decisions have smaller effects on final war outcomes than do choices at the strategic and grand strategic levels. An example might be German attempts in 1944 to estimate whether the Allied invasion of Europe would come at Normandy or Calais. Various constraints effectively ruled out sites either further west or east, but this still left the Allies (and thus the Germans) with a choice to make between the two remaining options. Had the Germans guessed correctly, the
Allies’ 1944–45 campaign would have been noticeably more expensive, but it is quite unlikely that the outcome of the war would have changed or that its length would have been affected very much. By comparison, strategic and grand strategic choices, which can exert larger effects on war outcomes, are usually heavily constrained by factors such as geography; states’ political, social, and material resource endowments; and so forth. Therefore strategic and grand strategic choices are often more similar to the mutual optimization model discussed above than they are to guessing games. For example, in World War II the Allies could not defeat Germany without moving huge amounts of men and matériel from the United States to Europe, and their only practical method was by sea. The Germans’ most efficient method of opposing this flow was by submarine. Both sides’ plans took these facts into account, but this was a mutual optimization problem, not a problem of guessing the other’s intentions. Thus, although strategic interaction can create situations where our best measure of the offense-defense balance becomes a band rather than a point, it is not clear that these bands are often wide. This is, however, a worthwhile avenue for empirical research.

Beyond these points that are closely connected to the concept of structure, Goddard and Finel address some additional important points. First, Goddard argues that in adopting a broad definition of the offense-defense balance by including nationalism and force size, we cannot separate the balance from power, because these factors influence power. In fact, nationalism can affect both power and the offense-defense balance, in different ways. As Goddard notes, nationalism can augment a state’s power by increasing its ability to extract resources. However, nationalism also affects extraction capabilities differentially, increasing them more when the state is trying to protect territory that is understood to be part of the national homeland than when it seeks to take territory that is not part of this homeland. Nationalism makes it easier to translate aggregate power into the ability to defend the national group’s homeland. It makes it harder to translate power into the ability to conquer territory that is perceived to be outside the national homeland. This differential effect cannot be incorporated into our standard notions of power; rather, it is best understood as a shift in the offense-defense balance in favor of defense.

Goddard is correct that force size can reflect power, but it should also be included among the factors that influence the balance. Some states simply lack the resources to deploy forces of the size that would create an offense-defense balance best matched to their goals. States then do the best they can with the resources available, that is, within the constraints imposed by their power, which determines the size of their forces. For example, the impact of conventional force size on the balance results from an interaction between power and geography. High force-to-space ratios tend to favor defense more than do low force-to-space ratios. As a result, two states whose mobilization potential is relatively large in relation to length of front will face a balance more in favor of defense than if both were weaker. For example, even if technology had remained constant, the balance between Germany and France in 1914 would have been more favorable to defense than it was in 1870. Both states were able to mobilize greater resources, which shifted the offense-defense balance as well as the balance of power. If we consider only power, we would predict the wrong impact on the outcome of war. Although from the 1870s onward Germany’s power grew faster than France’s, its chances of conquering France declined. We could have lumped this effect under the
general heading of “geography,” but we consider it important enough in its own right to mention separately.12

A main theme of Finel’s letter is that the offense-defense balance cannot be measured: the factors that influence the balance cannot be reliably coded, and the results of war games are not reproducible. We agree with Finel that whether the balance can be measured is central to the utility of the theory. In our article we explain, first, why the tools of military net assessment are well matched to measuring the balance.13 Also, although we offer reasons for optimism, we stress that our article does not settle the question of how closely the balance can be measured. In fact, the article concludes by calling for research into whether net assessment has been feasible in the past and whether net assessment techniques can be further improved. Second, our article explains how offense-defense theory can often make useful predictions even when net assessment is difficult and therefore substantial uncertainty about the value of the balance exists. Thus Finel’s criticism would be telling only if reliable net assessment can never be done. To accept Finel’s pessimism would mean, for example, that during the Cold War we did not and could not have had any real idea of Soviet prospects for a successful offensive on the central front, and that the German military before World War I lacked the analytic resources to detect the flaws and uncertainties in the Schlieffen Plan. If Finel holds these views, he could have contributed to the debate by spelling out why we should accept them.

Finel also objects to our version of offense-defense theory because he says it wrongly treats seizure of territory as the principal method of victory in war. This would indeed be a mistake, and would skew net assessments and thus estimates of the offense-defense balance—but this is not our position. Finel has simply misread us, confusing the measure of success with the means for achieving success. For purposes of the theory, we treat changes in political control of territory as the measure of success in war outcomes (i.e., we use the standard Clausewitzian definition). A military offensive should be considered a success if, at the end of the war, political control has changed to the advantage of the attacker—either because the attacker has gained full or partial control of territory it did not control before, or because it has undermined an opponent’s control of territory.14 The two most important means of achieving military victory are (1) destruction of enemy forces or mobilization potential, underpinning their ability to contest territorial control; and (2) credible threats of such severe punishment that the opponent would rather concede the territory than continue the conflict.15 It is true, however, that

12. Concern over nuclear proliferation illustrates a similar interaction between power and the offense-defense balance. Opponents of proliferation argue that new nuclear states will be unable to build the large forces required to provide adequate retaliatory capabilities; as a result, although nuclear weapons provided the superpowers with highly effective deterrent capabilities, they will not do the same for small countries with limited resources. In other words, according to this argument, the offense-defense balance for these less powerful countries is more favorable to offense than was the case for the superpowers because these weaker countries will have smaller forces.
15. There is some evidence that, at least in conventional conflicts, punishment is less likely to be decisive than is destruction of forces. Robert A. Pape, Bombing to Win: Air Power and Coercion in War (Ithaca, N.Y.: Cornell University Press, 1996).
who controls certain territories during a war can sometimes affect victory to the extent that the territory can be used to mobilize additional capabilities with which to pursue one or both of these approaches. Thus the offense-defense balance, which is defined as the cost of the means of victory relative to the cost of the means of denying victory, reflects the impact of territory only when it influences war outcomes.

Finally, Finel argues that even if the offense-defense balance could be adequately specified and measured, the cost in terms of complexity would outweigh the insights that offense-defense theory can provide. We agree that measuring the offense-defense balance is complex. Nevertheless, three points suggest that Finel is again too negative. First, as we argue in our article, key variables in other structural theories of international relations, such as power in balance-of-power theory, also become difficult to operationalize if we take the task seriously. This is simply a general problem in international relations theory. Second, as we discuss in our article, in cases where complexity makes complete net assessment intractable, analysts can often simplify their task while preserving reasonable confidence in their estimate of the balance. One possibility is to focus on particular theaters or campaigns that are expected to have a decisive impact on the overall war outcome. For example, in the 1980s, net assessments of a conventional World War III in Europe concentrated on estimating the prospects of success of the Warsaw Pact’s initial offensive into Western Europe. Given NATO’s much larger mobilization potential and control of the oceans, the Warsaw Pact could not hope to win a war if the initial offensive was not successful. Third, and perhaps most important, the predictions made by offense-defense theory often diverge significantly from standard power-based structural theories, which means that if the offense-defense balance can be measured, the payoff is likely to be worth the effort.

In closing, we continue to believe that our article has placed offense-defense theory on a firmer foundation. Further research is warranted, because offense-defense theorists have established powerful deductive arguments showing that power alone is insufficient to explain state decisions about military forces, strategy, and war, and that offense-defense variables should influence these decisions. It is too early to render a verdict from empirical testing of offense-defense theory, which is still in its infancy. Further testing is the key.

—Charles L. Glaser
Chicago, Illinois

—Chaim Kaufmann
Bethlehem, Pennsylvania

17. We discuss other types of simplifications in our article; see ibid., p. 61.