



CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

APRIL 2026

Methods of National Power Analysis: Pitfalls and Best Practices

Nicholas Kitchen

The Future of American Power Series

Methods of National Power Analysis: Pitfalls and Best Practices

Nicholas Kitchen

April 2026

About the Carnegie Endowment for International Peace

In a complex, changing, and increasingly contested world, the Carnegie Endowment generates strategic ideas, supports diplomacy, and trains the next generation of international scholar-practitioners to help countries and institutions take on the most difficult global problems and advance peace.

With a global network of more than 200 scholars across twenty countries, Carnegie is renowned for its independent analysis of major global problems and understanding of regional contexts.

© 2026 Carnegie Endowment for International Peace.
All rights reserved.

Carnegie does not take institutional positions on public policy issues; the views represented herein are those of the author(s) and do not necessarily reflect the views of Carnegie, its staff, or its trustees.

No part of this publication may be reproduced or transmitted in any form or by any means without permission in writing from the Carnegie Endowment for International Peace. Please direct inquiries to:

Carnegie Endowment for International Peace
Publications Department
1779 Massachusetts Avenue NW
Washington, DC 20036
P: + 1 202 483 7600
F: + 1 202 483 1840
CarnegieEndowment.org

This publication can be downloaded at no cost at
CarnegieEndowment.org.

TABLE OF CONTENTS

Executive Summary	1
Introduction	3
Power Analysis: Basic Conceptual Foundations	4
Power Measurement: Capabilities, Metrics, and Aggregation	6
Practicalities of Power Measurement by Domain	9
Conclusion	11
Notes	13



About the Author

Nicholas Kitchen is an Associate Professor in International Relations at the University of Surrey and the Executive Director of the Centre for the Study of Global Power Competition.

The Future of American Power Series

What is the future of American power? The United States commands extraordinary resources across every dimension of national power, yet in recent years it has struggled to achieve many important foreign policy aims. This paradox raises a key question for Americans and the world: What can the United States actually do with the power it has? Our project examines not just the quantity but the qualities of American power, assesses the degree to which it is eroding, and asks how the United States might use the power it has to better effect.

The Future of American Power series is edited by Christopher S. Chivvis, director of the American Statecraft Program.

American Statecraft Program

The American Statecraft Program develops and advances ideas for a more disciplined U.S. foreign policy aligned with American values and cognizant of the limits of American power in a more competitive world.

Executive Summary

Power is the foundational concept in political science and in particular International Relations (IR), yet the multidimensional nature of the concept means operationalizing it for the purposes of power analysis is fraught with methodological difficulty.

As a disciplinary undertaking, IR has used proxy aggregates of national power such as the Composite Index of National Capability (CINC) in order to explain patterns in the interactions of major powers over time. This modelling of “the balance of power” is a theoretical artifice that uses best-fit correlate variables to demonstrate broad structural patterns of balancing behavior.

On the one hand, these endeavors have had only mixed success when judged against the historical record. On the other, they have been misunderstood and misused as a future-facing analytical tool for foreign policy analysis, representing an aggregate national ability of a state to impose its will. This is a category mistake: power is multidimensional, and the exercise of power always takes place within a social relationship between actors. The notion of a single aggregate quantity of national power existing in the abstract therefore makes little or no sense.

Programs of national power assessment, then, face a conceptually difficult task. Since the dimensions and domains of power are not fungible or aggregable, and subjectivity and context are inherent to how power relationships operate, generating granular and detailed knowledge of any given state’s ability to apply resources in order to secure its objectives threatens to overwhelm analysts with myriad variables.

Furthermore, the selection of capabilities that have the disposition to produce power will depend on the purposes for which power is intended. Yet what it means to be powerful is subjective and contested: there is variance not simply in terms of the goals of policy, but at a more fundamental level of what constitutes the successful operation of power.

But some form of power assessment is crucial to governments, both as a matter of short-term policy formulation and as the basis for long-term planning. Strategic competition is not a single theater: there are different spheres of competition, each with their own logics of interaction, and tools of coercion and persuasion.

Disaggregating questions of national power into domains is therefore a useful first step for analysts seeking to evaluate the status of some broader notional “great power competition.” Within each, assessing power-position is not simply a case of understanding the relevant capabilities that may be productive of strategic advantage. Quality of data is crucial, but often absent, leading to reliance on proxy measures that do not directly assess capabilities themselves.

Volume assessments of capabilities need to be contextualized according to the dynamics of how they might be employed, taking into account structural features of the domain that determine how advantage is constituted. These might include factors such as geography, path dependencies, network and knowledge effects, and questions of whether what matters is power rank, power gap, or distance from the frontier. It is not simply place, but place within the distribution—that is, whether power is concentrated or dispersed—that determines power position.

Fine-grained, data-contextualized approaches to power at the domain level are a crucial part of a process of problem-definition and policy setting for states. However, this analytical task is rarely—if ever—rigorously done. Instead, there is a temptation to produce dashboards based on weighted proxy metrics that provide a false impression of clarity, outsourcing the phenomenon we're looking to understand to an easily accessible substitute. When such efforts provide us with both a score and justification for action, their outcome risks becoming the end of policy itself.

For this reason, strategic net assessment approaches have exhibited a skepticism toward data-driven approaches and eschewed any kind of integrated approach to national power analysis. Instead, net assessments have favored the use of diffuse, disconnected studies as part of a continuing diagnostic effort to understand the nature of the environments within which different aspects of competition take place, and how they may change.

The two approaches are not mutually exclusive. Qualitative domain assessments are a prerequisite for understanding which data might be relevant to gather, and how that should be contextualized analytically. The aim of power analysis should therefore not be to shortcut complexity, but to produce an impressionistic picture of the multiple domains and dimensions of power that embraces nuance, opening up conversations about long-term strategy rather than seeking to provide distinct answers to questions of immediate policy.

Power assessments shape our perceptions of the limits of the possible. Quantitative rankings and dashboards can provide false confidence, or focus our attention on the metric intended to reflect power, rather than power itself. Deeper, critical assessments, based on simulation methods, will derive less clear conclusions, but in doing so can open up policy conversations about resource investments, strategies of resilience, and policies of cooperation with allies and partners.

Introduction

Attention to power assessment among practical policymakers reflects the “return” of great power competition, in which major states view international affairs as fundamentally characterized by strategic competition between great powers. If states are operating in an era of great power competition, it is incumbent on them to know who is winning.

There are, however, different understandings of what winning means and how to recognize it. One view would be that it is enough to secure advantage in the metrics commonly used to measure power and have others accept that. Another view is that latent superiority of this kind is not enough, and compliant displays of obedience secured through the application of coercive capabilities are needed. Yet another group sees the resort to such displays of power as evidence of failure; they see power in the ability to shape consensual outcomes through influence. Still others see the need for these kind of continually active strategies as evidence of a lack of the most fundamental form of power: the existence of shared preferences cultivated through the manufacture of consent.¹

These differences may be superimposed onto the grand strategic posture of recent American administrations, from the capabilities focus of the 1992 Defence Planning Guidance, to the positive-sum dynamics of Clintonian enlargement, and the post-9/11 imperial assertiveness. The Obama administration claimed victories in the outcomes of painstaking multilateral diplomatic initiatives; in contrast, Donald Trump’s robust commitment to “winning” requires the creation of losers on the other side.

Such disagreements are important, because different ways of thinking have different—and sometimes contradictory—implications for the development of national strategy. They also have implications for how power is assessed. The purpose of this paper is therefore to explain pros and cons of different approaches to assessing national power through an explication of the premises behind the main approaches used today. While the conceptual and methodological considerations outlined here have bearing on the rigor with which any power assessment might be conducted, the effectiveness of such analyses will depend on their fit with an underlying vision of how power operates, and what national power is for.

To guide the reader through this reassessment, the paper proceeds in four stages. It first establishes the conceptual foundations of power as a context-specific capacity rather than a static collection of resources or data points. It then critiques the pitfalls of traditional aggregate measurement and proxy metrics, proposing instead that foreign policy analysts and national security strategists disaggregate national power into functional domains. The third section details the practicalities of assessing these domains by establishing a causal theory of power to identify where structural advantages truly lie. Finally, the paper introduces simulation as a critical tool for testing the conversion of latent advantage into relational influence, concluding with a warning against allowing reductionist “dashboards” to substitute for rigorous strategic thought.

Power Analysis: Basic Conceptual Foundations

At one level, the definition of power is relatively simple: Robert Dahl's classic definition, drawing on Max Weber's understanding of *Macht* as "the opportunity to have one's will prevail within a social relationship, also against resistance,"² captures the intuition that power is one actor's ability to get another to do something they would rather not. However, every major conceptual study of power notes the lack of agreement of how to conceive of and define power, and other, similar terms such as influence, control, authority, persuasion, and coercion.³ The myriad forms of power described in the literature suggest that power is a fundamentally elusive concept.⁴

Power Is a Capacity, Not an Outcome

Power is sometimes confused, or used interchangeably, with influence. Yet as Peter Morriss notes, the two terms—the former usually a noun, the latter more often used as a verb—express different ideas, and indeed, the two words have different etymological roots. "Power' always refers to a *capacity* to do things . . . an ability, capacity, or dispositional property."⁵ "Influence," while more diffuse in its usage, tends to refer to the (successful) exercise of that power.

This points us to the first of two fallacies of power analysis: to confuse power with its (successful) exercise. Here we meet a key methodological difficulty inherent in assessing power: to verify the existence of power, we must demonstrate ability through successful effects. Of course, being able to say who *was* powerful is clearly of less practical use than identifying who *is* powerful. But it is also not how we tend to use the concept: few would suggest that the United States was less powerful than North Vietnam in the 1960s. In a war that was asymmetric, the United States was not successful in achieving its goals, but that would not lead to the conclusion that it had been less powerful in that particular relationship or in a more general assessment. The *exercise fallacy* is the reason power should be regarded probabilistically, as the opportunity to prevail, as opposed to a demonstrated ability: power is not the same thing as the effects of power.

The dominant solution to the exercise fallacy in International Relations has been to think of power in terms of underlying resources: the "bases" of power, or more often, "capabilities." The phraseology of "power bases" points to a second fallacy, the *vehicle fallacy*, that power is not the same thing as the vehicle—or bases, or resources—that produce it. Capabilities—from diplomats and soldiers to money and guns—are not powerful in and of themselves: the likelihood that they can successfully produce power effects depends on the nature of the interaction within which they are deployed.

That is not to say that capabilities have no intrinsic potential. Particular resources, by their nature, possess the *capacity* to produce power effects in particular circumstances; that is their

disposition.⁶ For example, a bottle of whiskey on a shelf has the capacity to cause drunkenness. But it can only produce that effect when consumed, and its impact will depend on both the drinker (their weight and experience of consuming alcohol) and the context in which it is drunk (alone or with friends, or in a relaxed or lively environment). Such capacity is not a guarantee, but a probability: as Lukes puts it, “power is a potentiality, not an actuality – indeed, a potentiality that may never be actualized.”⁷

Power Is Context-Specific

Understanding power probabilistically requires that we understand the context within which it operates, including the structures within which interactions take place, the identity and aims of the actors, and the presence of risks and uncertainties.

Structural Context

Historical path dependencies, infrastructures, institutions, and ideas can enable and constrain the interaction of actors, creating “structural power” for some of these actors.⁸ Structural power may derive from relational advantage: at moments of system-making that follow major wars, dominant states may seek to create enduring arrangements that confer advantage into the future,⁹ which would generate costs should others seek to abandon those structures.¹⁰ Meanwhile, deep, sociocultural power emerges and shapes norms, perceptions, and preferences of an apparent natural order of things.¹¹ This “third face” of power leaves actors either unable or unwilling to conceive of alternatives.¹² Whether or not we ascribe this natural order to the “insidious” exercise of hegemonic power, as Lukes puts it, or conceive of it more benignly, it is easier for actors to succeed when their interests and preferences are aligned with the prevailing order.

Risks and Uncertainties

While capabilities may be disposed to produce effects, that probability is associated with risk, and subject to uncertainty. Calculable risks include the risk of misjudging power relationships, imperfect strategy implementation, and the other side’s response. Uncertainties include accidents, oversights, agency, and innovations. And while probabilistic analyses of state power tend to assume calculable risk, in practice, decisions to deploy power engage myriad risks and uncertainties, because of the multiplicity of interacting variables.¹³ Agile actors that have the *protean power* to adapt, improvise, and innovate more quickly are often better placed to cope with the unexpected.¹⁴

Power Has No Permanent Hierarchy

It has been an axiom in realist thought that the *ultima ratio* of power in IR is war, and that therefore military capacity, and the economic and other resources that underpin it, are the key determinants of strategic competition in the past and today.¹⁵ Yet while military power

is certainly *important*, in general, superior military power is not *decisive* either in armed conflict or in “peaceful” strategic competition. Since different forms of power operate in different ways, it makes little sense to speak of a permanent hierarchy of power resources.¹⁶ And while disputes may escalate in the direction of violence, power does not rest solely on the capacity for violence. The historical absence of nuclear-armed powers employing significant military force against one another, despite becoming engaged in significant disputes, places further doubt on notions of a linear hierarchy of power.¹⁷

Actors shape their power relationships by how they perceive their interactions: power can only emerge in a relationship when the actors psychologically permit it to do so.¹⁸ Weber’s concept of *Macht* points also to the distinct concept of *resistance*, against which the opportunity of realizing power capacity rests. Those in weaker power positions can raise the costs for more powerful adversaries,¹⁹ but they may also resist power on a deeper psychological and societal level, generating profound resilience that is not quantifiable.²⁰

The absence of a hierarchy of power resources, coupled with the capacity for actors to innovate and psychologically resist the apparent imperatives of power, should be cautionary for power analysts. Not only are these dynamics difficult to identify, predict, and measure, they can have a decisive impact on power relationships. The facts of power are always more fundamentally contingent and uncertain than they appear.

Power Measurement: Capabilities, Metrics, and Aggregation

The dominant method of power measurement in IR scholarship has been focused on measuring material capabilities in order to assess the bases of power. This epistemologically positivist approach rests on two presumptions, neither of which, as the previous sections have shown, are necessarily clear: first, that we can assume a relatively consistent relationship between latent power resources and outcomes of power relationships; and second, that power resources are measurable facts. Proceeding on this basis to the question of national power leaves us with two key methodological questions: which capabilities should we measure, and how should we go about measuring them?

Limits of Capability Metrics and Proxy Measures

There are inherent challenges in measuring material capabilities. Some power resources are more easily measured than others. For example, the destructive capacity of artillery can be quantified. Some latent power resources—land area, population, proven reserves of natural resources, steel production, to name a few—can also be judged with quantitative statistics.

Other common measures, however, are more problematic. For example, it is commonly agreed that significant research and development (R&D) capacity is an important underpinning base of power, but linking R&D to particular capabilities is much harder to do.

To get around challenges like this, proxy measures are often used, but these shortcut the requirement to conduct fine-grained assessments in ways that can mislead. For example, spending—one of the most common proxy measures—is a measure of input not output: cost overruns on capital programs, increasing veterans’ pensions costs, and large-scale corruption in procurement can all increase the proxy measure, while adding nothing to actual capability. Similarly, a focus on current spending ignores the impact of stocks of capabilities previously purchased.

Perhaps the most ubiquitous proxy measure of “national power” is GDP, which, despite its prevalence, fails to deduct welfare costs and relies on statistical sampling methods that are empirically impressionistic at best. Consider that, in 2014, a periodic rebasing of Nigeria’s GDP statistics saw its GDP grow by 89 percent overnight.²¹ Though no serious scholar of International Relations would conclude that Nigeria had become almost twice as powerful, that would be the conclusion of any number of the national power formulas.

These limits should give us caution in any quest for complete, quantitative measures of national power.²² Employing economic and systems analysis approaches can have perverse effects both in terms of military behavior and strategic decisionmaking, as during Robert McNamara’s tenure as secretary of defense, when enemy body counts and kill ratios were employed as a key performance metric in the prosecution of the Vietnam War.²³ Due analytic caution, however, does not mean we should reject quantitative measures. Statistical measures can present accurate pictures of specific capabilities and underpinning resources. The challenge is to understand what those pictures mean in terms of power, which requires care in how the results are contextualized, understood, and used. Selecting metrics is never a scientifically neutral act: the choices will shape who and what may be thought powerful.

Limits of “National Power” as a Concept for Practical Policy

In the discipline of International Relations, measures of power have been devised and operationalized for a particular theoretical purpose: discovering the balance of power. At the center of balance of power thinking is the idea of aggregate measures of national power, which create the distributions that inform balancing behavior, structure the international system, and explain the persistence of conflict. Aggregating national power, however, is more problematic than these approaches suggest. Aggregated measures of power select and weight variables in an essentially arbitrary way,²⁴ and are reliant on subjective “smell test” verification.²⁵ This indefinite underpinning leads to mismatched outcomes. When policymakers consider what tools they need to be successful in world politics, they are asking a somewhat different question to social scientists who are seeking to explain the causal determinants of major war. For social scientists, aggregate power is a useful baseline for theory building. But for practical policy purposes, the notion of “national power” has real drawbacks.

This is in part because which specific capabilities will be relevant depends on the nature of the power interaction in question, and international strategy and diplomacy introduces multidimensional problems that may be addressed in a variety of ways. But it is also a more fundamental conceptual problem: there is simply no standard of measurement by which various power resources can be assessed against one another and thus aggregated at the national level. Different dimensions of power and types of capability are incommensurable.²⁶ This would not be such a major issue if capabilities were fungible, that is, easily convertible from one form into another. Equally, if there was a single issue that dominated international politics, to which all power was intended to apply, aggregation might be possible by weighting the impact of capabilities against that standard. But neither is the case: the majority of resources exhibit very limited convertibility and produce specific and different power effects (which are context dependent), and nations develop broad and varied capabilities because the purposes to which they will be put are broad and varied. For these reasons, analyses that claim to produce a single aggregate measurement of national power do so only by making arbitrary choices of what to count and subjective decisions about how to weight them, and so risk misleading those who use them.²⁷

Domains and Metrics

The issues of fungibility and aggregation become less salient the more limited the scope of assessment. Focusing on distinct instruments or elements is therefore an established method for assessing capabilities. Formulations found in defense policy literature, such as diplomatic, informational, military, economic (DIME) or expanded variants (such as military informational diplomatic financial intelligence economic law development, or MIDFIELD) may appear to be exercises in acronym construction, but they do help to capture the different types of power assets available to states actors.²⁸

To be effective, however, an “elements of power approach” requires some degree of contextualization about the purposes to which capabilities are to be put, and the actors with which interaction takes place. Absent such context, policymakers may be encouraged to simply build-up in ways that increase the value of the instruments but do not necessarily increase the opportunity to prevail. In short, the size of a capability may be mistreated as inherently valuable even when it is not relevant, or may even be detrimental, to solving the policy problem at hand.

Thinking of power in terms of domains—the aspect of another state’s behavior you are seeking to affect (that is, the issue at hand)—is a more effective approach.²⁹ It allows for assessments to integrate the capabilities most relevant to such interactions. When defining domains, care should be taken to ensure that the aggregation of any indicators reflects relatively homogenous and substitutable phenomena. As a result, within-domain aggregation has been most effective surrounding economic, financial, and technological questions, where components can be assigned monetary value relatively unproblematically.³⁰ Elsewhere, in the military domain, for example, it is not obvious that naval, land, sea, space, and cyber assets can be integrated into a single measure without some clearly defined probabilistic assessment of the expected nature of military interactions.

Practicalities of Power Measurement by Domain

Assessing power within domains helps tether capabilities to the aspects of behavior to which they apply. But measuring power within a domain is not simply a question of generating quantitative volume measures of relevant capability. In order to assess power within a domain, we need to understand both the nature and operation of power within that domain. A number of considerations apply:

1. What is the causal theory of power?

A theory of power describes the causal chain between the data being assembled and the potential power effects of the capability. It has two main elements: first, what is the link between what is being measured (the metric, or indicator of capability) and the power resource itself? Essentially, how good is the proxy at capturing the power base we're trying to measure? Second, what is the nature of the capacity of that resource to generate power effects? How likely is it to produce what kind of effects, under what conditions, at what cost, and with which associated risks?

2. What is the nature of advantage?

How is power constituted in this domain? Are actors' capabilities purely relative, or is there a winner-take-all dynamic in which the dominant state derives a disproportionate benefit? Is advantage best understood hierarchically as a question of rank, or is it the distance between parties that matters, or distance from the leader? In what ways does geography matter? Are there institutional enablers or barriers, normative constraints, or other social constructs shaping the way in which power resources can be employed? Are there network effects, knowledge monopolies, or chokepoints that may enable particularly effective strategies for certain actors, or produce specific vulnerabilities for others?

3. What is the appropriate statistical measure?

Quantitative data can present an effective picture of power resources where it is operationalized within domains that can be fully captured by a single metric, or where domains exhibit enough homogeneity for fungibility and substitutability to enable statistical aggregation of component indicators. Determining the appropriate measure may be a simple question of whether nominal volume measure, global share, per capita, or otherwise adjusted or normalized data, is the most appropriate metric. Or it may be a more complex question of deriving an impressionistic composite from a series of proxy measures, each one thought to be a necessary component but none on their own sufficient.

4. Is there good data?

Economic studies of power have been tempted to work backwards from the available data without constructing a proper theory of power and have neglected dimensions of power that cannot be easily quantified, for example, the quality of diplomacy that is central to agenda-setting power. Data should, as far as possible, quantify the power resource itself, and not be a proxy for it.

Concentration and Dispersion

When using metrics to establish a picture of the operation and structure of power within a domain, and distribution of capabilities between actors, the resultant picture may demonstrate both concentrations and dispersion of power.

“Concentration” refers to a situation where an actor (or relatively few actors) has the capacity to bring significant weight to bear, either across a broad scope of actors, or in relation to the most significant actors. Concentration may derive from:

- a significant capabilities gap in terms of quantity (for example, the United States possesses more military aircraft than the next five nations combined);
- a qualitative differentiation in capacity as a result of technology (China’s dominance of ultra-high-voltage electricity transmission); or
- where structural or network dynamics create particular advantages (the centrality of the U.S. dollar and U.S.-regulated financial institutions to global payment systems).

“Dispersion” refers to the opposite situation, where the distribution of capabilities and the impact of structural arrangements create a situation in which no actor is particularly differentiated in terms of their ability to affect actors at scope. This does not mean that power is equally distributed—and both concentration and dispersion will appear different from the perspective of the relevant scope of actors. Dispersion is more likely where institutional arrangements mitigate divergences in capabilities or create positive-sum dynamics.

From Structural Advantage to Relational Influence

Assessments of concentration or dispersion within domains can indicate where structural advantage or vulnerability lies. To interpret that finding, however, it will be helpful to spell out key assumptions about the object and timeframe of competition within that domain, the limits of conduct, and the identities of the other actors. Converting a structural advantage within a domain into relational influence on the behavior of an actor may be as simple as pointing out the existence of the advantage. Alternatively, it may require the deployment of capabilities, augmented by that structural advantage, in order to coerce or incentivize.

It is at this point that *simulating* power relationships is an essential tool of power analysis. Simulation is primarily focused on understanding why actors are likely to behave in particular ways. Simulating power relationships allows analysts to examine, in a specified relationship of power competition, what capabilities, in what combination, when, where, how, and at what cost, actors are prepared to employ to secure their interests and achieve their goals. Simulations move us beyond power in principle—a divergence in capabilities, or an advantageous network position—to power in practice, enabling us to understand how actors

perceive the application of power and respond to it. Simulations are particularly useful when iterations show the commitment of actors over time: a more powerful but less committed actor may be able to prevail in the first instance, but may then become distracted, whereas the more interested actor will repeatedly return to the issue.

A variety of analytical tools, both formal and informal, fit under the banner of “simulation,” including military war-gaming, red-teaming, and stress-testing, in which the real-world process under analysis is imitated in order to be studied. Simulation is a qualitative technique, in contrast to formal modelling or systems analysis, which employ fixed variables and consistent, measurable criteria. What simulations allow is for inferences and evaluation to be drawn from not only the results of the exercise itself but also the insights and reflections of the participants.

Simulations are particularly useful for allowing policymakers to test the actual leverage in practice that a particular position of power provides. Decisionmakers regularly misperceive not necessarily their position, but what they can do with it, assuming that capabilities will produce expected effects and underestimating the capacity for innovation and will of the opponent. Interdependence, issue-linkage, blowback consequences, solidarity, and resistance all regularly produce unanticipated costs associated with the conversion of structural advantage into relational influence.

Moreover, the realization that structural advantage in a particular domain may be converted into coercive applications may incentivize efforts to alter the structural dynamics of the domain itself. This has been a concern of the so-called weaponization of interdependence, particularly in relation to the United States’ conversion of its structurally advantageous position in foreign currency and financial networks into highly coercive secondary sanctions, prompting calls for alternative payments systems and reserve currency instruments.

Conclusion

Questions of national power analysis tempt us toward a vision of the world where one country or another is “number 1.” The who’s up, and who’s down, akin to sports league tables, are attractive shorthands, capturing the public’s attention and enabling political actors to shape narratives. As Joseph Nye put it, “periods of ‘declinism’ tell us more about popular psychology than about geopolitics.”³¹

But power assessment will also shape perceptions of the limits of the possible. The probabilistic nature of national power assessment creates risks that those interpreting the results may not appreciate the contingencies inherent in the findings. Data-driven approaches in particular can provide false confidence: metrics change our understanding of value by outsourcing what we’re looking for to an easily accessible substitute, just as a FitBit reduces our

multidimensional goal of health and fitness to a singular metric of step counts, providing us with a score and with it a motivational scheme. With a FitBit, we can get clear feedback of how we're doing, in competition with ourselves and others. By outsourcing the process of value deliberation, we can stop thinking about what it really means to be healthy and just concentrate on how many steps we can do.³²

Proxy assessments of capability provide us with just such a shortcut. The goal of power analysis, however, should be deeper, to engage in questions of what power means and how it operates within domains. Building up a necessarily impressionistic picture that captures the multiple dimensions of power will not provide clear conclusions, but it can open up policy conversations about resource investments, strategies of resilience, and policies of cooperation with allies and partners. The temptation to create reductionist rankings, visualizations, or dashboards that strip the data of its essential context may result in misinterpretation by their users, even when disclaimers are included. Iterative use of such tools may result in their becoming the end of policy in itself, as decisionmakers surmise that a stronger position with respect to adversaries and competitors derives from policies that result in an improvement in the metrics being employed. Policymakers should be wary of concentrating on the indicator, rather than embracing the complexity of the social reality it purports to represent.

Notes

- 1 Adam Quinn and Nicholas Kitchen, “Understanding American Power: Conceptual Clarity, Strategic Priorities, and the Decline Debate,” *Global Policy* 10, no. 1 (2018): <https://doi.org/10.1111/1758-5899.12609>.
- 2 Felix Berenskoetter, “Thinking About Power,” in *Power in World Politics*, ed. Felix Berenskoetter and M.J. Williams (Routledge, 2007).
- 3 Harold D. Lasswell and Abraham Kaplan, *Power and Society: A Framework for Political Inquiry* (Yale University Press: 1950); Robert Dahl, “The Concept of Power,” *Behavioral Science* 2, no. 3 (1957): 201–15, <https://doi.org/10.1002/bs.3830020303>; Peter Morriss, *Power: A Philosophical Analysis* (Manchester University Press, 2002); Michael Barnett and Raymond Duvall, “Power in International Politics,” *International Organization* 59, no. 1 (2005): 39–75; Steven Lukes, *Power: A Radical View* (Palgrave Macmillan: 2005); Stefano Guzzini, *Power, Realism, and Constructivism* (Routledge, 2013); David Baldwin, *Power and International Relations: A Conceptual Approach* (Princeton University Press, 2016); Muhammad Aiman Nasuha Azari et. al, “The Three Dimensions of China’s Comprehensive National Power (CNP): Hard, Soft & Smart,” *Journal of Administrative Science* 19, no. 2 (2022): 23–38.
- 4 Kitchen notes non-exhaustively that “power has three dimensions, and four faces; it can be hard, soft, sharp, sticky, social, relational, structural, discursive, reciprocal, compulsory, productive, active, inactive, intentional, unintentional, extensive, intensive, disciplinary, authoritative, direct, indirect, potential, actual, distributive, zero-sum, institutional, diffused, symbolic, visible, invisible, hidden, public, private, ubiquitous, normative, civilian, usable, protean, latent, agenda-setting; it has an architecture, networks, communities, elements, levels, spaces, capillaries, boundaries, balances, and imbalances.” See Nicholas Kitchen, “Making Net Assessment Work: Evaluating Great-Power Competition,” *Survival* 66 (2024): 51–70. See also Mark Haugaard, “Power: a ‘Family Resemblance’ Concept,” *European Journal of Cultural Studies* 13, no. 4 (2010), 10.1177/1367549410377152.
- 5 Peter Morriss, *Power: A Philosophical Analysis* (Manchester University Press, 2002), 12–13.
- 6 Peter Morriss, *Power: A Philosophical Analysis* (Manchester University Press, 2002), 12–13.
- 7 Steven Lukes, *Power: A Radical View* (Palgrave Macmillan, 2005), 69.
- 8 Nicholas Kitchen and Michael Cox, “Power, Structural Power, and American Decline,” *Cambridge Review of International Affairs* 32, no. 6 (2019), 734–52: doi:10.1080/09557571.2019.1606158.
- 9 John G. Ikenberry, *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars* (Princeton University Press, 2001).
- 10 Nicholas Kitchen and Simon Curtis, “‘If You Build It, They Will Come’ Infrastructure, Hegemonic Transition, and Peaceful Change,” *Global Studies Quarterly* 5, no. 1 (2025): 4, <https://doi.org/10.1093/isagsq/ksaf021>.
- 11 Robert W. Cox and Timothy J. Sinclair, *Approaches to World Order* (Cambridge University Press, 1996), 151.
- 12 Steven Lukes, *Power: A Radical View* (Palgrave Macmillan, 2005).
- 13 Rose McDermott, *Risk-taking in International Politics: Prospect Theory in American Foreign Policy* (University of Michigan Press, 1998).
- 14 Peter J. Katzenstein and Lucia A. Seybert, *Protean Power: Exploring the Uncertain and Unexpected in World Politics* (Cambridge University Press, 2018).
- 15 Edward Hallett Carr and Michael Cox, *The Twenty Years Crisis, 1919-1939: An Introduction to the Study of International Relations* (Palgrave Macmillan, 2001); Robert Gilpin, *War and Change in World Politics* (Cambridge University Press, 1981); Kenneth Waltz, *Theory of International Politics* (Addison-Wesley Publishing Company, 1979).
- 16 Harold D. Lasswell and Abraham Kaplan, *Power and Society: A Framework for Political Inquiry* (Yale University Press, 1950), 94.
- 17 Carl Kaysen, “Is War Obsolete?: A Review Essay,” *International Security* 14, no. 4 (1990): 42–64, <https://doi.org/10.2307/2538750>; Michael Mandelbaum, “Is Major War Obsolete?,” *Survival: Global Politics and Strategy* 40, no. 4 (1998): 20–38, <https://doi.org/10.1093/survival/40.4.20>; John Mueller, “The Obsolescence of Major War,” *Security Dialogue* 21: 321–28; Nina Tannenwald, “The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use,” *International Organization* 53, no. 3 (1999): 433–68, <https://www.jstor.org/stable/2601286>.
- 18 Stefano Guzzini, “Embrace IR Anxieties (or, Morgenthau’s Approach to power, and the Challenge of Combining the Three Domains of IR Theorizing),” *International Studies Review* 22, no. 2 (2020): 277, <https://doi.org/10.1093/isr/viaa013>.

- 19 Anthony Giddens, *Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis* (University of California Press, 1979).
- 20 Jack Barbalet, "Power and Resistance," *British Journal of Sociology* 36, no. 4 (1985): 531–48, <https://doi.org/10.2307/590330>.
- 21 Diane Coyle, *GDP: A Brief but Affectionate History* (Princeton University Press, 2015).
- 22 Stefano Guzzini, "On the Measure of Power and the Power of Measure in International Relations," Danish Institute for International Studies, Working Paper no. 28, 2009, 5.
- 23 Matthew Fay, "Rationalizing McNamara's Legacy," *War on the Rocks*, August 5, 2016.
- 24 Frank Clifford German, "A Tentative Evaluation of World Power," *Journal of Conflict Resolution* 4, no. 1 (1960).
- 25 Richard J. Stoll, "State power, world views, and the major powers," in *Power in World Politics*, ed. Michael D. Ward (Lynne Rienner Pub, 1989), 135–57; J.P. Cole, *Geography of World Affairs* (Butterworths, 1983), 228.
- 26 Stefano Guzzini, "Fungibility of Power Resources," in *Encyclopedia of Power*, ed. Keith Dowding (SAGE Publications Ltd, 2011).
- 27 David A. Baldwin, *Paradoxes of Power* (Blackwell Pub, 1989), 167; Frederick W. Frey, "Comment: On Issues and Nonissues in the Study of Power," *American Political Science Review* 65, no. 4 (1971): 1087, <https://doi.org/10.2307/1953499>; Susan Strange, *States and Markets* (Bloomsbury Publishing, 2015), 11.
- 28 For a discussion of the development of instruments of power approaches in U.S. military doctrine, see Cesar Augusto Rodriguez et al., "Putting the 'FIL' into 'DIME': Growing Joint Understanding of the Instruments of Power," *Joint Force Quarterly* 97 (2020): <https://ndupress.ndu.edu/Media/News/News-Article-View/Article/2106566/putting-the-fil-into-dime-growing-joint-understanding-of-the-instruments-of-pow/>.
- 29 This presentation of domain represents natural usage, where scope refers to the extent of actors affected, in terms of number and importance. However, in the power analysis literature, domain and scope are sometimes used in the opposite way, such that scope refers to the issue and domain the extent. See, for example, David Baldwin, *Power and International Relations: A Conceptual Approach* (Princeton University Press, 2016).
- 30 It should be noted that, although we use them every day, the meaningfulness of bedrock concepts of macroeconomic analysis such as output and investment are often questioned precisely because they aggregate very heterogenous phenomena, simultaneously reducing questions of quality to questions of monetary value.
- 31 Joseph S. Nye Jr., "American Greatness and Decline," Project Syndicate, February 1, 2024, <https://www.project-syndicate.org/commentary/wit-trump-american-decline-becomes-self-fulfilling-prophecy-by-joseph-s-nye-2024-02>.
- 32 C. Thi Nguyen, "Value Capture," *Journal of Ethics & Social Philosophy* 27, no. 3 (2024), 469.



CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

1779 Massachusetts Avenue NW
Washington, DC 20036

[CarnegieEndowment.org](https://www.CarnegieEndowment.org)