



A production of H-Diplo with the journals *Security Studies*, *International Security*, *Journal of Strategic Studies*, and the International Studies Association's Security Studies Section (ISSF).

<http://issforum.org>

H-Diplo/ISSF Editors: **James McAllister and Diane Labrosse**

H-Diplo/ISSF Web and Production Editor: **George Fujii**

Commissioned for H-Diplo/ISSF by **James McAllister**

Response to H-Diplo/ISSF Forum on “What We Talk About When We Talk About Nuclear Weapons.” <http://issforum.org/ISSF/PDF/ISSF-Forum-2.pdf>

Published by H-Diplo/ISSF on **18 June 2014**

URL: <http://issforum.org/ISSF/PDF/ISSF-Forum-2-Response.pdf>

What We Do, and Why it Matters: A Response to FKS¹

Francis J. Gavin

“Don’t let us forget that the causes of human actions are usually immeasurably more complex and varied than our subsequent explanations of them.”

- Fyodor Dostoevsky, *The Idiot*

In his recent Jack Ruina Nuclear Age lecture at MIT, Robert Jervis – arguably our most important scholar of nuclear dynamics – reminded his audience how little we actually know about the influence of nuclear weapons. “Their impact on world politics is hard to discern.” Everywhere one looks, Jervis pointed out, there are puzzles that remain stubbornly immune to definitive answers. Would the Cold War have happened at all without nuclear weapons, or would it have unfolded in much the same way? Do nuclear weapons stabilize international relations or make the world more dangerous? Why don’t more countries have nuclear weapons? Why did American decision-makers pursue strategies and deployments that seem to have disregarded the fundamental insights scholars had proposed about the meaning of the nuclear revolution? Why is this gap even larger when you look beyond the United States to the eight other nuclear-weapons states?

¹ I am grateful to Mark Bell, Eliza Gheorghe, Nick Miller, Vipin Narang, Reid Pauly, and Jane Vaynman for their helpful insights.

Were scholars prescribing when they thought they were describing? Did the nuclear balance matter, and if so, when and in what ways? Were all conflicts between nuclear states in some sense nuclear wars? What role did credibility play in nuclear politics, given that deterrence is based on a threat to use nuclear weapons few actually believed? Perhaps most importantly, how have our *ideas* about nuclear weapons changed over time, and how have these changes affected the *realities* of nuclear weapons? Jervis's remarkable meditation was a pointed reminder that we lack certainty on these issues, and must be humble in our efforts to understand these terrifying, horrific weapons. The great challenge for scholars is "to recapture the strangeness of the nuclear world."²

I want to thank Matthew Fuhrmann, Matthew Kroenig, and Todd Sescher (FKS) for their thoughtful reply and willingness to engage in such an important subject in this novel platform. I'd also like to thank those scholars who commented on this debate, Scott Sagan for expertly framing these issues in his introduction, and in particular, the extraordinary H-Diplo team of George Fujii, Diane Labrosse, and James McAllister for shepherding a spirited discussion. I've have learned quite a bit from these exchanges – if nothing else, I am now far more conversant in terms like omitted variable bias and selection effects than I was a year ago. I hope others in the H-Diplo/ISSF community will now join the discussion.

Obviously, I disagree with much of what FKS say in their response. But a detailed reply to a reply to a review of two articles would be a bit silly. I think each of us has laid out his arguments clearly, and I leave it to others to weigh in and decide for themselves. I would, however, like to make two points: first, about how we should think about the role of methodology in our scholarship, and second, why these questions and debates around nuclear statecraft are of fundamental importance for both security studies and policy.

I do not propose – as FKS imply -- that the "most important example" method is the only or always the best way of explaining how the world works.³ If I were studying the links between smoking and cancer, where the "average" effects were important, I might well use quantitative methods -- though I would hopefully employ an N several orders of magnitude larger than theirs of 52 or 210 cases. I would additionally recognize that smokers may have been smoking different things in different quantities and that their smoking interacted with countless other distinct variables over each smoker's life to determine health outcomes. I might note that the first scientist to demonstrate the link between tobacco and cancer initially used *experimental* methods.⁴ Furthermore, I would

² Robert Jervis, "Why We Should Be Puzzled About Nuclear Weapons," comments delivered to the MIT Security Studies Program *Jack Ruina Nuclear Age Dinner*, March 4, 2014, Hotel Marlowe, Cambridge, MA.

³ Rather, I argued that the inability of their theories to explain far and away the most important case should give one pause, especially if you believe FKS's N's are inflated with cases where their theories are irrelevant. Political scientist often talk about easy, hard, and critical cases for testing their theories – it would be difficult to imagine a case that should be easier for them to prove or more critical to their efforts than the standoff between the Soviet Union and the United States between 1958 and 1962.

⁴ Robert P.N. and Angel H. Roffo, "The forgotten father of experimental tobacco carcinogenesis," *Bulletin of the World Health Organization* 2006; 84: 494–6. I might also puzzle over the disturbing (and long

acknowledge the recent concern within the biomedical community about many of its health findings based on statistical studies.⁵ Most importantly, I would recognize that different methods are appropriate for different questions, and insist that we should choose a method(s) based on how well it explains what we are trying to understand.

The most important point of my critique was to demonstrate the inability of FKS's method to explain (or even properly specify) the most dangerous nuclear crisis in history, the 1958-1962 standoff between the Soviet Union and the United States. As a result, I found little comfort that they claimed to explain outcomes in far less important confrontations where it is not even clear that nuclear possession played any role, like crises over Haiti or El Salvador. Given the rarity, difficulty of defining, complexity, and potentially horrific consequences of any nuclear crises, this is one subject where I am not as interested in "averages." In fact what I am really worried about and want to better understand when it comes to nuclear crises is "outliers." Others might feel differently.

FKS are of course right that statistics have many virtues, and international relations scholars have used them effectively to shed important light on many questions, including nuclear dynamics.⁶ Scholars often posit theories with observable implications and it is natural that we should want to test them as rigorously as possible. But caution and humility are in order. We often comfort ourselves with the belief that through math we can drain our analysis of prejudice, bias, ad-hockery, and other "unscientific" thoughts that many believe plague qualitative tests and narrative accounts. Dig deeper, however, and it

buried) historical fact that the earliest and most sophisticated epidemiological studies of the tobacco-cancer link were sponsored by Nazi Germany [Robert N. Proctor, *The Nazi War on Cancer* Princeton (NJ): Princeton University Press; 1999]. Perhaps I would use archival research to better understand what cigarette makers knew about the links between smoking and cancer, and when they knew it. K. Michael Cummings, Anthony Brown and Richard O'Connor, "The Cigarette Controversy," *Cancer Epidemiol Biomarkers Prev* June 2007 16; 1070.

⁵ John P. A. Ioannidis "Why Most Published Research Findings Are False," *PLOS – Medicine*, August 30, 2005, found at <http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0020124>. For a recent journalistic account of the failure of large-scale biomedical studies to produce consistent, replicable results and determine causality, see George Johnson, "An Apple a Day, and Other Myths," *The New York Times*, April 21, 2014, accessed at <http://www.nytimes.com/2014/04/22/science/an-apple-a-day-and-other-myths.html> For evidence of a similar crisis emerging in the quantitative social sciences, see Jerry Adler, "The Reformation: Can Social Scientists Save Themselves?" April 28, 2014, *Pacific Standard: The Science of Society*, found at <http://www.psmag.com/navigation/health-and-behavior/can-social-scientists-save-themselves-human-behavior-78858/> It should be noted that the biomedical studies under question use statistical methods far more sophisticated, have far larger and cleaner "N's", and are more easily replicated than those of FKS.

⁶ For an exemplary model that is both multi-method and takes the important selection effects problem head on, see Nicholas L. Miller, "The Secret Success of Nonproliferation Sanctions," Forthcoming, International Organization, Fall 2014. For an excellent paper that deals with the issues raised by FKS, see Jonathan Renshon, Vipin Narang, Arthur Spirling, and Jane Vaynman, "Fool's Gold: The Role of Nuclear Weapons in International Conflict," Prepared for the 2014 Annual Convention of the International Studies Association, Toronto, Ontario, March 2014.

becomes obvious that those numbers often reflect similar untested assumptions, biases, and interpretations, and are no more scientific than qualitative accounts.⁷ Consider a simple but important question that goes to the heart of the issue: what *is* a nuclear crisis (or *when* does a crisis become nuclear)? I imagine that we could generate as many answers to this key question as there are subscribers to H-Diplo. Or how would we define, let alone quantify, an “observation” (FKS seem to use “case” and “observation” interchangeably, which is confusing) -- is the whole crisis an observation, or does it consist of a series of observations, given the constant ebb and flow of strategic interaction, gaining information, and learning between adversaries?⁸ Needless to say, international politics does not take place in the stable, clean environment of a lab. FKS’s analogy of the zero-sum, closed system of a sporting match obviously fails to capture the complexity of nuclear crises, where identifying the winner and the loser of a nuclear standoff is often in the eye of the beholder and can change over time.

The real problem with this kind of analysis is that statistics can be very powerful tools to establish correlations, but often more problematic – especially as they are used by FKS – to establish causality.⁹ The divorce rate in Maine, for example, correlates precisely over time with the per capita use of margarine in the United States, and the changing rate of people killed by their own bedsheets maps almost exactly with shift in U.S. ski facility profits, but no one would seriously argue for any causal inference in these examples.¹⁰ Large-n

⁷ In fact, they can be worse, since what might be thought of as the “hexing powers of science” and the methodological bullying that often takes place in the academy can have a chilling effect on debate and discussion.

⁸ In the same way that central bank monetary policy and antibiotics do not have the same effect over time as household/firms and bacteria learn and adapt, one can imagine leaders pursuing different policies over different points during a crisis based on “anticipatory adaptation” and interactive learning. In other words, even if the military balance remained the same, United States policy during the Cuban Missile Crisis may have ended up much differently if the Kennedy administration had been forced to make a decision on October 16th, 1962, as opposed to having had almost two weeks to learn, interact, and adapt, coming up with a policy by October 27/28 few would have proposed earlier. On monetary policy and antibiotics, see John H. Makin, “Endogeneity: Why policy and antibiotics fail,” January 30th, 2014, *Outlook*, American Enterprise Institute, <http://www.aei.org/outlook/economics/monetary-policy/federal-reserve/endogeneity-why-policy-and-antibiotics-fail/>

⁹ Philip A. Schrodt, “The Seven Deadly Sins of Contemporary Quantitative Political Analysis,” *Journal of Peace Research* 2014 51: 287. For popular accounts highlighting the overselling of statistics and big data, see Tim Harford, “Big data: are we making a big mistake?” *FT Magazine*, March 28, 2014, accessed at <http://www.ft.com/intl/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz2xa1MyJfW>; Gary Marcus and Ernest Davis, “Eight (No, Nine!) Problems With Big Data,” *The New York Times*, April 6, 2014, accessed at http://www.nytimes.com/2014/04/07/opinion/eight-no-nine-problems-with-big-data.html?smid=fb-share&_r=0 Needless to say, the data sets of FKS are *far* smaller and less homogenous than typical “big data” data sets, making these problems even more pronounced.

¹⁰ “Funny Graphs show correlation between completely unrelated stats,” May 9th, 2014, found at <http://twentytwowords.com/funny-graphs-show-correlation-between-completely-unrelated-stats-9-pictures/> For a more serious example of a failed effort to use big data to make correlations that ultimately were unconnected to causality, see the background behind the Google Flu Trends failure; Stephen Salzberg, “Why Google Flu is a Failure,” *Forbes*, March 23, 2014, available at

observational analyses in which selection effects, endogeneity, post-treatment bias, model dependencies, reverse causality, and non-comparable and temporally unstable cases (to use the lingo of my political science friends) are rampant do not allow for powerful claims to causal inference.¹¹ Even when statistical correlations are identified, it often does not translate into useful policy predictions.¹² Nor does this method help us understand the crucial but unobservable crises that *never happened* because of selection effects; in other words, where nuclear dynamics deterred the state from provoking a crisis in the first place.¹³ If what we want to know is *why* something did – or did not -- happen, FKS's methods fall short.

It is important to note that this methodological critique is not simply a matter of “different horses for different courses.” Historians have been wrestling with statistics and “big data” for well over a century, and are quite aware of its benefits and shortcomings, most famously exposed during the rancorous debate over Robert Fogel and Stanley Engerman's use of statistics to make claims for the benefits of slavery in the pre-Civil War U.S. South.¹⁴ Reflecting on the great controversies within the discipline of history over what was dubbed “cliometrics,” the former President of the Organization of American Historians and the American Historical Association Joyce Appleby “observed that while quantitative history had made it impossible to deny the structural inequities in American history, the statistics had not spoken for themselves. Historical analysis required revealing the power relations that produced the numbers, and that work would generally be qualitative in nature.”¹⁵

<http://www.forbes.com/sites/stevensalzberg/2014/03/23/why-google-flu-is-a-failure/> For an analysis that says the real lesson from the Google Flu fiasco was better data, not more, see Kaiser Fung, “Google Flu Trends’ Failure Shows Good Data > Big Data,” Harvard Business Review Blog, March 25, 2014, available at <http://blogs.hbr.org/2014/03/google-flu-trends-failure-shows-good-data-big-data/>

¹¹ I am grateful to Vipin Narang for explaining these factors to me.

¹² Michael D. Ward, Brian D. Greenhill, and Kristin M. Bakke, “The perils of policy by p-value: Predicting civil conflicts,” *Journal of Peace Research* July 2010 vol. 47 no. 4 363-375. For an analysis of how widespread and troubling this issue is, see Regina Nuzzo, “Scientific method: Statistical errors,” *Nature*, February 12, 2014, found at <http://www.nature.com/news/scientific-method-statistical-errors-1.14700>

¹³ See the discussion of selection effects on pp. 25-26 of my original essay, <http://issforum.org/ISSF/PDF/ISSF-Forum-2.pdf>

¹⁴ For an excellent summary of how historians have thought about statistics and big data in the past, and innovative suggestions for how to exploit these tools for understanding foreign policy and international affairs in the future without making the kinds of mistakes that mar the FKS efforts, see David Allen and Matthew Connelly's unpublished paper, “Diplomatic History After the Big Bang: Using Computational Methods to Explore the Infinite Archive.”

¹⁵ *Ibid.*, p. 6. For Appleby's assessment, see Joyce Appleby, “The Power of History,” *American Historical Review* 103 (1998): 5-6. For a brief primer on the debate over Fogel and Engerman's historical work, see Nicholas Crafts, “Robert Fogel, controversial scholar who pioneered ‘cliometrics’”, June 16, 2013, ft.com, <http://www.ft.com/intl/cms/s/0/72555a02-d504-11e2-b4d7-00144feab7de.html#axzz305rjyA4k>

Do FKS actually believe their startling claim that qualitative methods are "particularly ill equipped for assessing causality?" (FKS, 11) Their previous work indicates they do not. As Todd Sescher wisely suggested in his 2007 dissertation: "Acknowledging the inherent limits of quantitative analyses, chapter 4 conducts a detailed case study in an effort to illustrate the mechanisms by which reputational factors might influence governmental decision-making during crises."¹⁶ Fuhrmann has also acknowledged the benefits of historical work for generating causal claims: "Given that every empirical approach has drawbacks, a multi-method assessment of my theory can inspire greater confidence in the findings presented in this article. The case study analysis above provides rich descriptions of my argument and illustrates that the causal processes operate as expected in actual instances of proliferation."¹⁷ In his book, *Exporting the Bomb*, Kroenig uses qualitative methods to "explore the mechanisms" of his theory by conducting an "in-depth analysis of an important case."¹⁸ I think Vipin Narang said it best in his contribution to this discussion:

"the causal inference revolution in quantitative methods may lead to a resurrection in the discipline's valuation of qualitative methods in nuclear security, since qualitative methods in this particular area are much better suited to identifying and teasing out causal mechanisms and processes than the big-data enterprise."¹⁹

This is not to say qualitative methods are without their own set of problems; no method is perfect. But I suspect many on this list would find FKS's claims about historical work puzzling. Given the volume and intensity of debate over important events in the past, historians might be surprised to learn that qualitative research "is not always so amenable to external oversight." (FKS, 7) We all understand that the historical record, just like the data sets that rely upon it, is often incomplete, and no historian I know would disagree with their claim that "the absence of evidence is not the evidence of absence." (FKS, 9). Historians deal with issues like incompleteness or participant bias in a variety of ways, including rigorously interrogating the evidence and engaging in multi-archival, multi-national research.²⁰ FKS and other like them should be especially grateful for this rigor, as

¹⁶ Todd Sechser, *Winning Without a Fight: Power, Reputation and Compellent Threats in International Crises*, PhD Dissertation, Stanford University, 2007.

¹⁷ Matthew Fuhrmann, "Spreading Temptation: Proliferation and Peaceful Nuclear Cooperation Agreements", *International Security*, volume 34, issue 1, summer 2009, p. 23.

¹⁸ Matthew Kroenig, *Exporting the Bomb: Technology Transfer and the Spread of Nuclear Weapons*, (Ithaca: Cornell University Press, 2010), p. 66.

¹⁹ Vipin Narang, "The Promise and Limits of Quantitative Methods in Nuclear Studies."

²⁰ It is the extraordinary increase in the number, quality, and accessibility of archival sources from around the world related to nuclear dynamics that is one of the most compelling argument for encouraging both historians and political scientists to mine these new sources. Wouldn't it be far better to encourage Ph.D. students to generate new knowledge, policy insights and answer questions that until now have been hidden behind a wall of secrecy, instead of using statistics to manipulate old data sets built on sources that are increasingly obsolete, incomplete, or flat out wrong?

this qualitative work is the basis for the inputs into their models, even if careless coding at times simplifies complex historical findings into blunt variables.

It is these very difficulties of actually knowing why something happened in world politics that make historians far more cautious about generalizations. Some of this, no doubt, reflects differences in temperament between the disciplines. One of our most distinguished diplomatic historians, for example, suggests that the whole concept of an “independent variable” is misleading at best.²¹ In the end, however, good historians and political scientists are similar in that they don’t simply let “the evidence speak for itself.” They base their insights on the constant interaction between their theories and conceptual frameworks and what they find in the empirical record.²² Though both sides may hate to admit it, what political scientist like FKS are trying to do when they “analyze” is not that different from what historians do when they “interpret.” As such, when we make bold causal claims and generalizations about important subjects, we should expect to have those analyses/interpretations, and the empirical base, methods, and assumptions that back them, exposed to rigorous examination. I invite the members of this listserv (and beyond) to undertake such an examination, both of my arguments and those of FKS.

Why does any of this matter? Isn’t this just so much academic posturing? That is certainly not how I see this debate and its significance. How and in what spirit we approach nuclear dynamics reveals what matters to us as scholars, and whether we can lay any claim to be taken seriously by people outside of the academy.

To understand the importance of these issues, both in the academic world and in policy, think of a brilliant student interested in understanding how the world works, and in particular, wanting to know how nuclear weapons influence international relations. Lets call her Isabel A. Isabel A recognizes she needs to learn more, and considers applying to graduate school to earn a Ph.D. and study with great professors. She first explores history departments – Isabel A majored in history as an undergraduate and assumed knowledge of the past would be good preparation for thinking about the future. But she is warned that no one in a top-ten ranked history department is interested in supporting a student, no matter how smart, who wants to study these dreadful weapons, particularly if she is interested in generating knowledge to help make better policy.²³ Next, Isabel A looks at

²¹ For this critique, see John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York: Oxford University Press, 2004). Gaddis also points out that statistics do a bad job of capturing the causal dynamics that often look less like a linear model and more like the punctuated equilibrium dynamics seen in evolutionary biology, like the sudden and unanticipated events that transformed Central Europe and the Soviet Union between 1989 and 1991.

²² For an excellent guide relevant to both historians and political scientists interested in international relations, see Marc Trachtenberg, *The Craft of International History: A Guide to Method*, (Princeton University Press, 2006).

²³ The steep decline in diplomatic and international history in American universities has been well noted.

political science programs. They at least appear to share her interest in nuclear dynamics, so she enrolls. But she finds herself spending most of her time taking methods classes, and instead of gaining substantive knowledge about the world, feels like she is being trained to become a mediocre statistician. Reviewing the top journals and the results of the academic job market, she notes that this discipline appears to reward, at least some of the time, methodological prowess over original insight about international relations.²⁴ Depressingly, Isabel A. finds few colleagues or mentors who encourage her curiosity and enthusiasm for important, real world questions. She is only thankful that the latest fad in the profession, natural experiments, cannot be applied to nuclear dynamics.²⁵

Now imagine Isabel A, years later. Fed up with the pathologies of the ivory tower, wanting to make a difference in the world, she becomes a national security official in the United States government. She advises a principal on incredibly complex and dangerous problems – what to do about Iran and North Korea’s nuclear program, how to react to Japan’s and

“Job openings on the nation’s college campuses are scarce, while bread-and-butter courses like the Origins of War and American Foreign Policy are dropping from history department postings In 1975, for example, three-quarters of college history departments employed at least one diplomatic historian; in 2005 fewer than half did.” Patricia Cohen, “Great Caesar’s Ghost! Are Traditional History Courses Vanishing?” *The New York Times*, June 10, 2009, <http://www.nytimes.com/2009/06/11/books/11hist.html?pagewanted=all&r=0> Even those positions that are labeled diplomatic history rarely are held by scholars who study great power politics or the influence of nuclear weapons on international relations. For an insightful piece highlighting these trends and lamenting the declining cooperation between historians and political scientists, see David Paul Nickles “Diplomatic History and the Political Science Wars”, *Perspectives on History*, May 2011, <http://www.historians.org/publications-and-directories/perspectives-on-history/may-2011/political-history-today/diplomatic-history-and-the-political-science-wars>

²⁴ Causing a renowned statistician to reveal his fears about the consequences of what and how he and his colleagues taught their brightest students. “I sometimes have a nightmare about Kepler. Suppose a few of us were transported back in time to the year 1600, and were invited by the Emperor Rudolph II to set up an Imperial Department of Statistics in the court at Prague. Despairing of those circular orbits, Kepler enrolls in our department. We teach him the general linear model, least squares, dummy variables, everything. He goes back to work, fits the best circular orbit for Mars by least squares, puts in a dummy variable for the exceptional observation, and publishes. And that’s the end, right there in Prague at the beginning of the 17th century.” David A. Freedman, “Statistics and the scientific method,” in W.M. Mason & S.E. Fienberg (Eds.), *Cohort analysis in social research: Beyond the identification problem* (New York: Springer-Verlag), 1985. I thank Reid Pauley for bringing this to my attention. For a recent “inside baseball” critique of rewarding shoddy statistical work in political science, see Schrodt, “The Seven Deadly Sins of Contemporary Quantitative Political Analysis.” This trend towards methods over substance has led to far less emphasis on theory and more on “hypothesis testing,” which has resulted in research questions being more narrow, less interesting, and of decreasing appeal to anyone outside of the political science discipline. See John J. Mearsheimer and Stephen M. Walt, “Leaving theory behind: Why simplistic hypothesis testing is bad for International Relations,” *European Journal of International Relations*, September 2013 vol. 19 no. 3 427-457, online copy available at <http://mearsheimer.uchicago.edu/pdfs/Leaving%20Theory%20Behind%20EJIR.pdf>

²⁵ A nuclear deterrence failure being one experiment we all hope is never run. For a very interesting and innovative article that uses experiments to test norms and measure attitudes towards nuclear use, see Daryl G. Press, Scott Sagan, and Benjamin A. Valentino, “Atomic Aversion: Experimental Evidence on Taboos, Traditions, and the Non-Use of Nuclear Weapons” in the *American Political Science Review* (February 2013).

Saudi Arabia's constant demand for reassurance, fears that the rivalry between India and Pakistan could spiral into a nuclear exchange.. Isabel A no longer possesses any tribal affiliations to a particular academic discipline or method – she is desperate for and will use any and all knowledge that can help her recommend the best policies and avoid catastrophe. Sadly, little of what the ivory tower offers is of use to her, and her policy colleagues are quite unimpressed with the cutting edge methods offered by her former discipline.²⁶ Their efforts to forecast are lamentably bad; time and time again, her former friends and colleagues from the social sciences offer theories and predictions that are proven wrong by real world events, with few consequences for their careers and almost no self-correction.²⁷ International relations scholarship does not seem to recognize that there are no dichotomous, easy choices in her world; instead, she is confronted by radical uncertainty, unintended consequences, and the theory of second best.²⁸ Looking over statistical studies of nuclear dynamics, she notes that even good estimates of average effects are of little help in making decisions, since they provide no insight into the causal mechanisms that are critical to understanding which policies to choose.²⁹ The one question

²⁶ “Aside from economics, the scholarly disciplines policymakers found of greatest interest were area studies and history....compared to other disciplines, political science did poorly.” This is especially true of non-qualitative approaches. “Conversely, the more sophisticated social science methods such as formal models, operations research, theoretical analysis, and quantitative analysis tended to be categorized more often as “not very useful” or “not useful at all,” calling into question the direct influence of these approaches to international relations.” See Paul C. Avery and Michael C. Desch, “What Do Policymakers Want From Us? Results of a Survey of Current and Former Senior National Security Decision Makers,” *International Studies Quarterly*, 2013, 1-20, accessed at <http://www.phibetaiota.net/wp-content/uploads/2014/01/Carnegie-Stimson-Article-On-SocSci-and-Policy.pdf> Despite the utility of history to policy, however, many of Isabel A.'s friends from history departments might condemn her decision to work with the U.S. government and view her as a sell-out.

²⁷ As Philip Tetlock famously demonstrated, experts are no better at forecasting the future of world politics than non-experts, and are often much worse. Professionally, experts – unlike decision-makers -- almost never suffer consequences for their bad predictions. See Philip Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton University Press, 2006); for a nice summary, see Louis Menand, “Everybody’s an Expert: Putting Predictions to the Test,” *The New Yorker*, December 5, 2005, available at http://www.newyorker.com/archive/2005/12/05/051205crbo_books1?currentPage=all Economists -- a group many political scientists want to emulate and even appear to envy – are even worse at forecasting, as their disastrous record in the period leading up to the 2008-09 financial crisis reveal. See Tim Harford, “An astonishing record – of complete failure :‘ In 2008, the consensus from forecasters was that not a single economy would fall into recession in 2009’” *The Financial Times*, May 30, 2014, available at <http://www.ft.com/intl/cms/s/2/14e323ee-e602-11e3-aeef-00144feabdc0.html#axzz33DAqOJ92>

²⁸ Francis J. Gavin and James B. Steinberg, “Mind the Gap: Why Policymakers and Scholars Ignore Each other, and What Can be Done About it?,” *Carnegie Reporter*, Spring 2012, available at <http://carnegie.org/publications/carnegie-reporter/single/view/article/item/308/>

²⁹ Though perhaps we should not idealize Isabel A's policy life. As an unnamed but sharp observer who knows both worlds well pointed out, a more realistic description of Isabel's life might be: “Isabel advises a principal on complex nuclear dynamics. Unfortunately, her boss already knows how those work, and asks her to find ways to support his views. Isabel spends most of her time writing talking points and clearing them three times over with twenty different offices, then the Secretary's office just changes them anyway. She hears that some friends in Policy Planning are interested in thinking about future emerging issues in nuclear proliferation, but who can get a job there? Desperate, Isabel considers becoming a Republican in order to land

she is desperate to know the answer to – why have nuclear weapons not been used since 1945, and what can keep that streak going – is rarely studied by her former community, since there is little incentive to pursue research where there is no variation on the dependent variable.³⁰

Is this too harsh an assessment of our disciplines? Perhaps. But surely we can do better than we have in recent times. As the wise namesake of the center I had the honor of being associated with, Bob Strauss, once said about foreign policy and national security, “This ain’t beanbag we’re playing. These are big-time issues, this is life or death, this is the future of nations.”³¹ Understanding nuclear weapons and their influence on world politics is too important, too consequential, to be driven by academic trendiness, methodological preening, or narrow disciplinary concerns. Scholars should be honest about both the possibilities and limitations of our methods, and on the lookout for buried assumptions and even deeply hidden prejudices that affect our perspectives.³² Finally, we should keep Isabel A. in mind during this discussion. Her experience wrestling with these issues likely makes her sympathetic to Jervis’s poignant reminder that these questions are as difficult as they are important, and willing to embrace his call for humility. As we continue to discuss and debate these critical questions in an honest, rigorous, and transparent manner, so should we.

© 2014 H-Net: Humanities and Social Sciences Online



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

a political appointee slot in the next election. Or going back to grad school." For a humorous take on this process as it relates to think tanks, see Jeremy Shapiro "Who Influences Whom? Reflections on U.S. Government Outreach to Think Tanks" *Brookings*, June 4th, 2014, http://www.brookings.edu/blogs/up-front/posts/2014/06/04-us-government-outreach-think-tanks-shapiro?utm_source=Sailthru&utm_medium=email&utm_term=%2ASituation%20Report&utm_campaign=2014_Situation%20Report

³⁰ For the difficulty but necessary task of exploring the history of what hasn't happened – a thermonuclear war – see Francis J. Gavin, "How Dangerous? History and Nuclear Alarmism" in *A Dangerous World? Threat Perception and U.S. National Security*, eds. John Mueller and Christopher Preble (Washington, DC: Cato Institute, 2014).

³¹ Robert S. Strauss, last United States Ambassador to the Soviet Union and first to Russia, in testimony before the U.S. Senate Foreign Relations Committee, quoted in Terry Atlas and Timothy J. McNulty, "Nixon Offers A Lesson for Bush," *Chicago Tribune*, March 12, 1992.

³² For a remarkably thoughtful meditation on methods and diversity that is relevant to both quantitative and qualitative approaches, see Christopher Achen, "Why Do We Need Diversity in the Political Methodology Society," April 30th, 2014, *The Political Methodologist*, <http://thepoliticalmethodologist.com/2014/04/30/we-dont-just-teach-statistics-we-teach-students/>