30-Jun-2020  
  
Dear Mr. Gannon:  
  
We have now received two very thorough reviews of your manuscript (see below). Both reviewers highlight parts of the paper that they find interesting and worthy of further development. Unfortunately, both reviewers also point out numerous weeknesses and ultimately recommend rejection.  
  
This does not provide us with a sufficient basis for further consideration of your manuscript for publication at Security Studies. We hope that the very detailed comments provided by each reviewer will help you as you continue to develop your argument and revise the paper for submission to another journal.  
  
We wish you luck in this endeavor.  
  
Sincerely,  
Prof. Randall Schweller  
Editor in Chief, Security Studies  
[schweller.2@osu.edu](mailto:schweller.2@osu.edu)  
  
  
Reviewer(s)' Comments to Author:  
  
**Reviewer: 1**  
  
Comments to the Author  
While this paper makes some important and interesting notions, overall I think it is not focused enough and at current stage its contribution – theoretical and empirical – is too limited.  
Gray zone  
First, I have some problems with the focus on gray zones. This is mainly because as used by the author this is a very broad concept, which is used to explore a too wide variety of cases. In fact, I had various additional comments on this issue during reading the first sections of the paper, but the more I read it I became more convinced that the author needs, in my opinion, to reframe the paper and to focus on cyber deterrence.  
Such a focus has a number of advantages: it allows to omit the long discussion on limited war; it allows to make the empirical discussion more focused (as the main case studies concern only cyber-attacks), and it removes the ambiguity regarding some arguments such as those presented in the tables (e.g., according to the author a gray zone is based on a dyad and not only on the challenger, but this is not reflected in the tables). Furthermore, I think, as I discuss below, developing the discussion on cyber deterrence in the context of the empirical cases may lead to important findings.  
Theoretical discussion  
I feel that the theoretical discussion (as well as the methodological discussion regarding the variables and categories used in the paper) is not as strong as it could be. This leads to some difficulties in establishing the empirical arguments. For example, while the author acknowledges that geography is not the variable itself that explains the variation, it is used too many times. This is not a problem per-se, but it obscures other useful explanations some scholars have considerably developed within deterrence scholarship, such as the notion of balance of interests. This concept can aptly explain the empirical cases. The point is that the theoretical section provides a limited guidance to really address the explained phenomenon.   
In a similar way, while the author provides some notions about resolve which seems to be crucial for the author's arguments, the paper can further develop this discussion in various ways, especially given the debate over deterrence credibility that the author alludes to (p. 18). While the author mentions a few important works in this regard the books of Kertzer (2016, Princeton) on Resolve in international relations, of Lupovici (2016, CUP) on wars of resolve, and of Lupton on Reputation for Resolve (2020, Cornell) are useful in providing a more nuanced view of whether and how resolve/credibility/reputation influence the strategy of deterrence.   
Empirical discussion  
I find the empirical discussion too speculative and the empirical evidence as not established enough. While this is obviously a challenge given the type of cases, and especially given the focus on cases of cyber deterrence, I think that this cannot be enough to support the statements the author makes, or enough to examine the theoretical assertions.  
For example, the author argues that "Our theory predicts that a more forceful Western response would have only escalated the situation since Russia’s actions were chosen through a calculation that its objectives could be accomplished at reasonable cost." (pp. 31-32). However, this is an example of how the selection of cases is problematic. The cases do not include cases that could distinguish between different levels of resolve of Russia.  
Likewise, the author notes that "Russia would probably lose a conventional contest with NATO, risking nuclear escalation in the process. Russia acts circumspectly as a result". Maybe, but this argument should be established on some evidence/references. Furthermore, a key question is not whether NATO will win, but rather whether NATO is willing to be involved in such a war with Russia (and for what goal, e.g., protecting Ukraine).  
  
Most importantly, eventually the cases and how they are presented and discussed show variation, but not causality or enough attempts to establish it. In this respect, the author uses the concept of Western deterrence, but it is not clear what Western deterrence is. Is it the deterrent strategy of the EU? of NATO? One could also wonder whether each of these actors employ this strategy, how and how it is evident? In any case, this requires further elaboration and looking into the practice of this strategy—for example by focusing on issuing threats. In other words, the strategy of deterrence cannot be assumed.  
In a similar way, a key question is how Russia sees these cases, how it perceived the resolve of its rivals, and specifically how it responded to it. The author only focuses on what happened-- the (observable) activity of Russia, but I'm not sure that in this context, deterrence success can be drawn from it. In other words, more empirical evidence is required.  
Cyber deterrence  
As noted above, I suggest re-structuring the paper in the context of cyber deterrence. This requires further elaborating on this scholarship and connecting it to various aspects that appear in the theoretical discussion such as resolve. Such a discussion will allow to compare more similar cases (as opposed to the discussion of the various different cases of gray zone, which include Chechnya (1999) on the one hand and the involvement in American elections on the other.  
That said, it should also be noted, that there are still significant differences among the four main cases of cyber attack.  
In addition, such a focus may help to develop interesting insights regarding cyber deterrence success. As the author notes, "deterrence gradient still matters in cyberspace, furthermore, we see Russia conducting low-intensity cyber influence and espionage operations around the world, while it conducts high-intensity cyber-physical operations in closer proximity to its border"  
I think this can be a really important contribution. But it requires as I noted above further empirical discussion as well as theoretical elaboration to locate these findings in the context of cyber deterrence scholarship and the debate over its effectiveness.  
Additional comments  
- The last section is too long.  
- The term deterrer (e.g., p. 18) is somewhat problematic since it implies that deterrence has worked.  
- On how deterrence success and failure is part of a continuum see Rid's (2012) CSP article.  
  
  
**Reviewer: 2**  
  
Comments to the Author  
To the author:  
  
I have, with regret, advised against publication.   Both the immediate topic – Russian “gray zone” behavior” – and the larger theoretical issues you explore are fascinating ones. Kudos on what you have accomplished.  
  
Let me walk through, below, the concerns I am left with.  My sincere apologies if I have misunderstood, or simply missed, any elements of your argument.  My comments are extensive not because I dislike the manuscript but rather for the opposite reason.  I bother to write at length because I recognize the potential.  Please go ahead and stick pins in a voodoo image of me:  though you have no way of knowing it, I am in fact a horrible person and doubtless deserve the pins.  But the extended comments are in fact meant to be constructive.  If you are a drinker, go ahead and pour yourself a drink before you start reading.  
  
1.  Let me start with my honest puzzlement regarding research design.  Explaining Russian behavior is in itself an important service.  Research on this topic is a valuable contribution to the needed knowledge base of both political scientists and practitioners.  If this were your goal, your research design, per se, makes sense (though many of my other concerns would still remain).  You suggest, however, that your manuscript is in fact aimed at advancing and defending a larger theoretical proposition – that “the scope and intensity of revisionist contests should… vary with the resolve of the revisionist and inversely with the credibility of deterrence” (p. 4) – and that the examination of Russian interventions is not an end in itself but a means to “test this proposition.”   In other words, in this article you are using a comparative-case methodology to test hypotheses derived from your theory against four cases of Russian behavior during the 1994-2017 period.  This approach encounters several problems.   
  
The first is the classic one for small-n research designs.  Do we have any reason to believe that findings from the Russian case are generalizable – for example, across time, or regime type, or level of technology, or level of interdependence, or system polarity?   Perhaps we do have good reason to think so – but there is a burden of proof, or at least a burden to demonstrate plausibility, that you rests on your shoulders, and I don’t see that you attempt to address this burden.   
  
The second is one of case selection.  Even if I simply assume that all revisionist powers under all conditions are essentially similar and there are reasons to assume generalizability across some larger universe, I am left wondering whether I should regard the Russian case-comparison as an easy test for your theory, a hard test for your theory, or some sort of critical test for your theory.  If the Russian case(s) fail to disconfirm your hypothesis (that is, if your findings “support” your theory), what are we to make of this?  Does this imply we should have confidence that it is true in other cases, or even in some identified subset of other cases?  
  
Third, if indeed the point of this exercise is to test your theoretical claim, I am also left puzzled by your decision to look at Russian cases.  It is very difficult to directly determine the value one of your critical independent variables (resolve).  We have no direct evidence about how high Russian resolve was in these cases.  We have to infer resolve from other information.  To a nearly equal degree, you are forced to try to  find some sort of proxy measure for Russian assessments of the credibility of U.S. deterrent threats.  We are guessing at the actual determinants of Russian decision-making based on very limited observed behavior.  More on this below, but this raises real concerns that you are in fact using your dependent variable to measure your independent variables.  
  
My point here is that unless there is some compelling reason to test your theory against Russian behavior in the 1994-2017 period that you haven’t told me about, I am absolutely stumped why you picked this/these case(s) to study.  Why are you picking a case on which you have so little data?  Why not look at cases on which we have very clear data about “resolve” and about the credibility (that is, the credibility of the deterrer’s threat in the mind of the revisionist power)?  Why not, for example, look at American behavior, about which we have a wealth of information?   As you say, Russian “gray-area” behavior isn’t unique.  Other great powers, including the United States, engage in gray-area behavior when they have revisionist objectives.  If we want to explain why we see this type of gray-area behavior by great powers, why not look at cases on which we have lots of information?  We can observe the same types of behavior exhibited by Russia toward Georgia, Ukraine, and Estonia in U.S. behavior toward (e.g.) Cuba, Nicaragua, Chile, (post-war) Italy, (early-Cold War) eastern Europe and Baltic states, and Iran (both in the Mossadeq period and in the Stuxnet virus case) – and, with the exception of Stuxnet, we have detailed, accurate, reliable information on the thinking that led to the decisions about the level of “revisionist” action and the impact of deterrent threats on behavior.  Even if you don’t want to look at the United States, why not look at some other more transparent actor, perhaps Britain or France?  
  
2.  One potential justification for publishing a piece of research, even if the research design leaves concerns, is that new data is presented.  Obviously, though, that is not the case here.  You are not presenting new information about Russian behavior or decision-making.  That’s ok.  Schelling did not do any original historical research either.  But it is a little surprising to read a manuscript explicitly on Russian behavior that is completely reliant on secondary or tertiary sources.  
  
3.  At least two of your cases pretty clearly do not fit into the universe of cases you claim to be examining.  
  
You use Mazarr’s definition of of gray-zone conflict as being “a carefully planned campaign operating in the space between traditional diplomacy and overt military aggression.”  (p. 9)  We can, I suppose, argue about whether it was Russia or Georgia that was the revisionist power in the Russo-Georgian War, but what is clear is that it did in fact involve overt military aggression.  This was a traditional  war.  It was not a gray-area activity.  It isn’t entirely clear that it was even a limited war, except in the sense that Georgia did not fight to the last soldier and that Russia was efficient in its use of military power.  It is an example of Russia being willing to go beyond gray-area behavior, not an example of Russian gray-area behavior.  
  
At the other end of your cases, the Russo-American case also does not fit within your defined universe.  You have indicated (page 3) that you are interested in cases in which neither side “is limited in its means (capabilities).”  You then lay out a gradient of means running from “info ops” to “conventional military (ground).”  (See, for example, Figure 2, on page 27.)  In the Russo-American case, however, the actual range of Russian capabilities is severely truncated.  It may be that Russia lacks the resolve to send tanks rolling down Pennsylvania Avenue toward the White House.  Never having sat in on a Kremlin meeting, I couldn’t say with confidence.  But what is certainly true that it lacks the capability to do so.   I suppose that one can engage in philosophical arguments about whether Russia has a “capability” to engage in “conventional military air/sea” operations against the United States, but I don’t know of any American military planner who thinks Russia has any serious capability to hurt the United States with purely conventional military air/sea operations.  I also don’t know of any American military planner who thinks that Russia has any serious ability to engage in “paramilitary” operations against the United States.  In other words, in terms of the gradient of activities discussed in Figure 2, Russian is limited by its capabilities (regardless of its resolve or its belief in U.S. deterrent threats) to no more than “cyber disruption.”  
  
In other words, of the four cases examined, only two – Estonia and Ukraine – seem to fit within the universe of cases that you are claiming to explain.  
  
4.  There is something of a larger point here.  Russian conventional kinetic capabilities are in fact severely limited by geography.  You posit that geography matters because it can be viewed as a proxy for deterrent credibility and/or for resolve.  Frankly, per below, I think a very strong prima facie case can be made that this is simply wrong.  But in any case, to the degree that images like the one in Figure 2 suggest that – as you say on page 28 – “overt military intervention occurs only in Russia’s immediate periphery (‘near abroad’),” I think you need to acknowledge that outside the “near abroad,” Europe, and some relatively easily accessible areas in the Middle East and Mediterranean basin, Russia does not possess any significant capability for overt military intervention.   In other words, with regard to conventional (and even paramilitary) military capabilities, Russia does indeed face a huge geographical gradient.  Reverting to Figure 2 again, for a moment, this figure, not surprisingly, shows no Russian conventional military interventions in Bolivia or Zambia.  One does not have to make arguments about Russian “resolve” or the credibility of American deterrents to explain this.  Were I a Russian military planner and told to drive one of my tanks down the streets of Lusaka or to bomb La Paz, I would have to point out that I was physically unable to carry out this order.  
  
To me this suggests that if you are going to try to explain why Russia uses only gray-area means in some cases, or why it does not use particular gray-area means (such as “little green men”) in some cases, you have to carefully and explicitly bound the cases you are going to consider.  
  
5.  Even given your decision to look at Russia as your revisionist power and to look at the four cases that you have chosen, I am puzzled by the fact that you seem to have assumed that Russian objectives were the same in all four interventions that you look at, and that behavioral differences must therefore reflect level of resolve in achieving these commitments or the deterrent threat.  In fact, the four interventions were (as far as we can tell, given our limited knowledge about Russian decision-making, as noted above) aimed at completely different objectives.  At least arguably, the fact that Russia employed different means in the four cases you are looking at can be explained simply by reference to what it was Russia sought to accomplish.  As far as we can tell, the Russian objective in Georgia was to force Georgia to acknowledge that Russia would have veto power over its future decisions.  The Russian post-Maidan objective in Ukraine was to seize a particular piece of territory (Crimea), to indefinitely destabilize the Ukrainian government, and to indefinitely hamstring Ukrainian economic development.  (In other words, having written off the possibility of reducing Ukraine to Belarus-like satellite status, the Kremlin decided simply to make sure that Ukraine remained a political and economic basket-case.)  Frankly, the Russian objective in Estonia remains entirely unclear.  Even post-hoc explanations are difficult to make convincingly.  The Russian objective in the United States was to try to achieve the election of a particular candidate and/or to undermine confidence in the democratic process. That different choices regarding means were made is hardly surprising, and it does not seem to require an appeal to “resolve” to predict the differences.  
  
6.  Given that one of your two independent variables is “deterrent credibility,” and what (except in the Russo-American case) the deterrent about which you are talking is an extended one, I worry that you seem to me to be inclined to conflate the target state and the deterring state.  For example, consider page 13 where you argue:  “gray zone conflict must be preferred by both sides in a contest; a target must also choose not to escalate.”  I think what you are intending to argue is that “the deterring power must also choose not to escalate.”  That is, when you say “both” I think you are referring to the revisionist actor and to the deterring actor, not to the revisionist actor and the target.  I think this is going to have some interesting wrinkles when it comes to the question of credible commitment.  
  
7.  This gets to the concept of “commitment.”  Although, as shorthand, we might say, for example, that “the United States is committed to Estonia,” what we are really saying is that the United States is committed to taking action X if Russia carries out action Y against Estonia.  Although, again as shorthand, one might say that “Russia doubts the credibility of U.S. commitment to Estonia,” often what one is really trying to say “Although Russia fully believes that the U.S. is committed to respond to action Z, Russia does not think that the U.S. is committed to respond to action W.”  In other words, it isn’t a question of credibility at all.  It is a question of whether or not Russia can achieve its objective (and, again, this is one reason why figuring out what Russia’s objective is in each of the cases you are examining is so important) without involving the United States.  
  
8.  I think, too, that you may want to think more – or at least help your readers think more – about your concept of “resolve.”   Often when the word “resolve” is used, it is meant in the context of “commitment.”  That is to say that if the United States is resolved to do such-and-so, it is committed to do such-and-so.  As I understand your argument, however, this is not at all what you mean by “resolved.”  What you mean by “resolved” is that an actor values an objective highly enough to fight a high intensity war to get it.  In your use of the word “resolved” it would, for example, be oxymoronic to say that the U.S. is resolved to use all means short of war to assist the Uigher people in their struggle against Chinese oppression.  Again, “resolve” does not exist as a general quality.  My understanding is that you have defined it as a yes/no variable:  does the actor highly value a particular outcome (where the “outcome” includes all of the things that happened along the way in getting to some final state), where “highly value” is measured by whether or not the actor believes the outcome is worth more than the other net costs of waging “a high intensity war” against a particular adversary in a particular location at a particular time.  This in turn involves knowing the probabilistic assessments of what “a high intensity war” would look like in particular situations (e.g., would the adversary escalate to nuclear weapons, would the adversary fight to the death, would the adversary enlist allies…..).   
  
9. As I’ve already noted, I find your argument that distance-from-the-United States can somehow be used as a proxy measure for credible commitment, or that geography can be used as a proxy measure for resolve, to be on its face implausible, given the historic record.  The United States, for example, seems to have been credibly committed to an anti-Soviet Turkey, but not credibly committed to an anti-Soviet Egypt or Iraq.  The United States was resolved to overthrow Ho Chi Minh, but it was not resolved to overthrowing Castro.  The United States appears to have been resolved to protecting France against a Soviet invasion long before it was committed to protecting Spain or Portugal against one….  Frankly, I think you either need to look at cases in which you can get data on resolution and perceptions of credible commitment directly or find some other proxy measure.  
  
10.  Honestly, I am afraid that I can’t make heads or tails out of your distinction between wars limited by ends, by means, and by risk.  You may have identified a very useful typology.  And perhaps I’m simply too old to think clearly any more.  But I can’t follow your argument.  And if I can’t – given how much time I’ve devoted to understanding distinctions like the one you’ve made – I suspect that others may also have a problem.  Why is a revisionist behavior like, say, U.S. post-World War II financial payoffs and disinformation in Italy to undermine the Italian Communists and to keep the Christian Democrats in power, best understood as a war limited by ends than as a  war limited by means?   Or when you say  (page 6) that “the Korean War exemplified an underappreciated type of war fought to achieve political ends short of traditional military victory despite having the capability to do so” are you suggesting that if the United States and China had engaged in an unlimited military war, using all of their military potential, the U.S. would have achieved, in any meaningful sense, its “political ends” in Korea?  (And if this is what you are arguing, what is it that you think America’s “political ends” were in Korea that would have been consistent with this?)  
  
11.  Though I am sort of able to follow your argument, I have some of the same problems figuring out exactly what you mean by efficiency and effectiveness.  At the bottom of page 15, when you actually offer a definition, you lose me entirely.  “Efficiency,” you say, “minimizes the cost and risk of a policy.”  Surely, for a revisionist, doing nothing would be the most efficient way to try to carry out a policy.  Admittedly, nothing is gained, but doing nothing would minimize cost and risk.  “Effectiveness,” you say, “maximizes influence over the outcome.”  What is “influence” over an outcome?   Do you mean “likelihood of achieving an outcome?”  If so, how do we define outcome?  I probably could maximize the likelihood of killing the Russian leadership by using enough nuclear weaponry on Russia to cause a nuclear winter, while simultaneously releasing every biological weapon I was conceivably able to produce.  Does this mean that this is the most “effective” policy?  I know it sounds like I’m just being argumentative, but I’m not.  I really don’t understand how you are defining these terms here.  
  
  
My apologies for having gone on at such length.  Again, let me emphasize that I am doing so not because I dislike your manuscript or because I am trying to be mean.  To the contrary, I wouldn’t spend my time on this if I didn’t think what you have proposed to do and what you have done were potentially so valuable.  I truly commend your effort.