THE PEARSON INSTITUTE
FOR THE STUDY AND RESOLUTION OF GLOBAL CONFLICTS

The Pearson Global Forum
Part II. Understanding Conflict Through Data
Deterrence in Cyber Space by Ethan Bueno de Mesquita, Sydney Stein Professor and Deputy Dean for Promotions and Recruitments, Harris School of Public Policy

ETHAN BUENO DE MESQUITA: A couple of years ago in the fall of 2016, I was lucky to be part of a small group of academics invited to spend a day in conversation with the senior leadership of the NSA and US Cyber Command. It's always interesting being at one of our intelligence agencies. You leave your electronic tethers behind, there's signs everywhere, warning that there's uncleared personnel in the area. You have conversations almost exclusively in acronyms.

But this visit was different. Although we didn't know it, those agencies at that time were confronting the most serious cyber attack in our country's history. Because the Russian sought to disrupt and undermine the presidential election. And the leaders that we were talking with, were in the midst of fighting and ultimately losing a political battle to be allowed to retaliate against Russian cyber aggression.

They were palpably angry in a way I had never before seen from senior military and intelligence leaders. And though they couldn't tell us exactly why, they made clear that the United States and its allies were under attack, and that while we had tremendous technological capacity, what we lacked was a strategy to effectively defend ourselves in cyberspace.

The last time the world faced a similarly transformed strategic landscape, was at the dawn of the nuclear age. And at that time, game theory played an important role in helping to formulate the key doctrines of nuclear strategy and especially nuclear deterrence.

I left that day of meetings shaken, but also motivated by the conviction the game theoretic tools might once again be of service to national security and peace. Along with two collaborators, I started reading, and learning, and thinking, and eventually writing about cyber strategy. And the ideas that I'm privileged to get to share with you all today, are the first modest results of those efforts.

That is not the right slide. Cyber Warfare is emphasized here by William Lynn. Is characterized by an attribution problem. Unlike in nuclear, or conventional warfare, in cyber warfare we are often uncertain who is responsible for an attack or even whether an attack has occurred at all. And this was the point of departure for me and my colleagues in thinking about deterrence in cyberspace.

There's an obvious sense in which attribution problems must weaken deterrence. Deterrence works by a threat. I tell Chris that if he attacks me, I'm going to hit him so hard, he'll regret having initiated hostilities. And if he believes me, he won't attack. But if there's an attribution problem, that is if there's some chance that he thinks when he attacks me, I'll mistakenly think it was Jim and I'll hit Jim instead, I'm going to have trouble deterring Chris.

And this is why, deterrence in cyberspace will never be as effective as it is in some other realms of warfare. But our analysis suggests that the effects of attribution problems on deterrence in cyberspace are deeper than this. Because of the attribution problem, deterrence in cyberspace is fundamentally multilateral. When you can see where the missiles are coming from, you can think in bilateral terms. Our nuclear second strike capability detours the Russians, and theirs detours us. The Chinese don't really come into it.

But when there's no return address, that bilateral logic is off the table. And to see why I think it's helpful, to think through an example. So consider the problem for the United States government in trying to assess blame, for the 2014 attacks on Sony pictures. In that blame assessment, one input were features of the attack itself.

Little snippets of code, they were similar to malicious code that had been used in earlier text, where the target was South Korea. IP addresses believed to be associated with elements of the North Korean Army and so on. But another input to the blame assessment, were more general features of this strategic environment and North Korea is placed within it. Imagine that some adversary in this case, the North Koreans, is believed to have become more capable, more aggressive in cyberspace. Then whenever a difficult to attribute attack occurs, that adversary is more suspect.

But if the North Koreans are more suspect, that must mean other adversaries, the Russians, the Iranians are less suspect following a difficult to attribute attack. And therefore less likely to face retaliation. Which of course makes it more tempting for them to in fact engage in cyber attacks. In effect, they can hide their activities, behind the activities of the already highly suspect North Koreans, which is precisely what the Russians attempted to do with their attacks on the PyeongChang Olympics.

And this is why attribution problems make deterrence in cyberspace fundamentally multilateral. If we become worse, at deterring any one of our adversaries, we become worse at deterring them all. And so in formulating a cyber strategy, we must think, and we must act globally because changing our strategy with respect to one adversary will change the behavior of all of our adversaries.

And for this reason, it's a little bit disappointing that the just released DoD Cyber Strategy which contains many important and I think productive strategic shifts, seems so single-mindedly focused on China and on Russia. Even if our primary strategic goal, even if our only strategic goal, is to deter the Chinese and the Russians, we must act to deter the North Koreans, and the Iranians, and Eisele and so on, because if we fail to deter those adversaries, we create a strategic environment that simply invites greater aggression from Russia and China.

Now, what does all of this imply for how we might think about a new deterrence doctrine for the cyber age, beyond just multi-lateralism? Traditional deterrence theory teaches us that we should seek to commit ourselves to greater retaliatory aggressiveness. The classic example of this, is the doctrine of mutually assured destruction, from nuclear strategy. We tell the Russians that if they attack us, we'll respond with such overwhelming force, we'll destroy their society, and they tell us likewise.

Now, once an initial Russian attack occurs, we might not want to go down that road. But if we've tied our hands that we will go down that road, we can deter the initial aggression. And so serious conversations are underway for how we might tie our hands, to commit to greater retaliatory aggressiveness across the Board in Cyberspace. Richard Clark, for example, has recently proposed that the United States hold governments responsible for all cyber attacks emanating from their domain, from their territory, regardless of who the perpetrator turns out to be.

But the analogy from traditional deterrence, to cyber deterrence on which such an analysis rest, is flawed. In conventional warfare, in nuclear warfare, the risk of retaliating against the wrong adversary is vanishingly small. Missiles come with a return address. But in cyberspace we face a fundamentally different trade off. Of course, committing to greater retaliatory aggressiveness, detours more attacks. But committing to a policy of retaliating more aggressively in cyberspace where attacks are often difficult to attribute, also means more frequent retaliation against the wrong adversary with attendant risks of dangerous Escalatory Spirals.

It also creates incentives for deliberate provocation by adversaries, rogue actors, looking to leverage the attribution problem to foment global conflict. And so, greater retaliatory aggressiveness across the board is unlikely to be optimal in cyberspace. We should commit ourselves through Doctrine, through Treaties, through Standing Military Orders, to greater retaliatory aggressiveness, following attacks that are particularly straightforward to attribute.

But having done so, we should also commit ourselves, to retaliate less aggressively than we would otherwise be inclined to do. Following attacks that are particularly difficult to attribute. Such forbearance will reduce the risk of erroneous retaliation, and reduce incentives for deliberate provocation at little cost in terms of deterrence, and security.

Let me end by saying, that deterrence is only the tip of the iceberg for cyber strategy. The cyber age has posed a host of deep strategic challenges. The interaction between cyber and conventional warfare, the potentially destabilizing effect of using cyber tools for tasks such as missile defense, the role of secrecy versus transparency in Cybersecurity, the implications of a conflict landscape in which responsibility for orphans and retaliation rests with the government, but where responsibility for defense and target hardening now resides in a diffuse marketplace and on and on.

Developing a strategy that comes to grips with these new challenges at the cyber age, is essential for the future of peace and prosperity. There's much work to be done.

Thank you.