

RED TEAMING SIMULATION EXERCISE

SUMMARY OF KEY FINDINGS

SEPTEMBER 2025



OVERVIEW

The Global Partnership Against the Spread of Weapons and Materials of Mass Destruction (GP) seeks to understand and counter state-sponsored or state-adjacent disinformation across the Chemical, Biological, Radiological, and Nuclear (CBRN) spectrum. As part of this effort, the Johns Hopkins Center for Health Security (JHCHS) conducted a **red teaming study** to assess the impact of Russian biological weapons (BW) disinformation on global non-proliferation, arms control, and disarmament norms. Findings can be used to inform efforts to safeguard these norms.

Red teaming is a methodology that involves adopting the persona of a decision maker and simulating that decision maker's thinking and actions. Used frequently by both academics and defense analysts, red teaming is considered a **rigorous methodology**, one that facilitates exploration of decision making while avoiding mirror imaging bias (when researchers project their world view onto a decision maker, resulting in incorrect assumptions and conclusions). In this exercise, participants were asked to adopt the role of a county leader and simulate that leader's decision making on issues related to BW pursuit. Findings from the exercise suggest that **disinformation can have real impacts on state-level decision making regarding BW**, underscoring the importance of addressing information manipulation in diplomatic settings to safeguard nonproliferation norms and cooperative threat reduction efforts. Indeed, we found evidence that **disinformation has the potential to increase state actors' interest in pursuing an offensive BW capability**, along with evidence that **debunking false allegations can be an effective strategy** to mitigate some of the consequences associated with disinformation.

METHODOLOGY

Participant recruitment and characteristics: To assess how different levels of expertise may impact state leaders' decision making, the following groups of participants were recruited to complete the exercise:

- 1) **40 BW experts** (individuals with expertise in BW-related concepts).
- 2) **40 country experts** (individuals with expertise in the country to which they were assigned)
- 3) **20 non-expert participants**

10 participants (4 BW experts, 4 country experts, and 2 non-experts) were assigned to each country. To ensure appropriate assignments and the validity of study outputs, participants were assigned to countries based on their level and area of expertise, both of which were assessed through an online questionnaire. 10 countries were included in the exercise. These countries were chosen to ensure a sample with representation from all three BWC regional groups and sufficient variation across political systems, alliances, and historical experiences with BW.

Prior to completing the exercise, participants were provided with two primers to ensure they had sufficient background knowledge to engage with exercise materials: one on biological weapons concepts, and another describing the history, domestic politics, and foreign policy of the country to which they were assigned (along with any relevant information about the country's history of BW pursuit and official stance on BW).

Exercise scenario: The red teaming exercise, which was adapted from a previous exercise,¹ was conducted through the online survey platform Qualtrics. The exercise scenario was based on real events, namely Russia's 2022 request for a BWC Article V Formal Consultative Meeting to discuss unfounded allegations of U.S.-funded BW development in Ukraine. Participants' interest in pursuing an offensive BW capability and/or a biodefense capability was then measured (both quantitatively and qualitatively) at three points in time:

- 1) **At the beginning of the exercise**, which was set in April 2021, prior to the Russian invasion of Ukraine.
- 2) **After the introduction of the first experimental inject (inject 1*)**, which was based on the disinformation presented by Russia during the 2022 Formal Consultative Meeting.
- 3) **After the introduction of the second experimental inject (inject 2*)**, which was based on the United States' and Ukraine's official responses to Russia at the 2022 Formal Consultative Meeting.

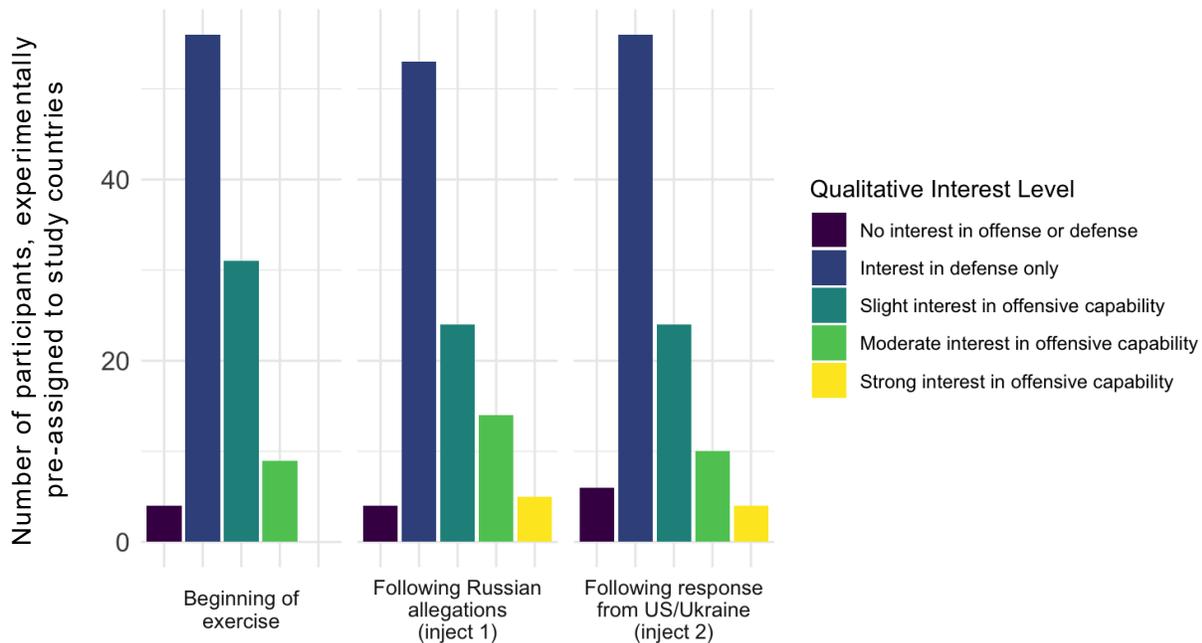
*Both experimental injects were shortened or edited (for clarity) versions of the original statements and documents presented at the Formal Consultative Meeting. No fictional claims or documents were added by the investigators.

Outcome measurements: The primary outcome (participants' interest, as leaders of their assigned country, in pursuing an offensive BW capability) was measured on two different scales:

- 1) **Written scale**, which required participants to select one of the following written statements:
 - 1) "I don't think our country should pursue any activity related to biological weapons, whether for offensive or defensive purposes." (*No interest in offense or defense*)
 - 2) "We should be conducting activities to enhance our defenses against biological weapons, including research to characterize deliberate biological threats and the development of medical countermeasures. However, we should not pursue offensive activities." (*Interest in defense only*)
 - 3) "That is a good question, and while we're not interested in pursuing an offensive biological weapons capability now, there are some circumstances in which an offensive program might theoretically be useful." (*Slight interest in an offensive capability*)
 - 4) "We should give careful thought as to whether acquiring an offensive biological weapons capability in the short to medium term is in our interest." (*Moderate interest in an offensive capability*)
 - 5) "An offensive biological weapons program is something that we ideally would like to pursue." (*Strong interest in an offensive capability*)
- 2) **Numeric scale**, which required participants to rate their interest in pursuing an offensive BW capability on a 0-10 scale (with 0 indicating no interest and 10 indicating strong interest).

Participants' interest in BW-related activities at each of the three measured time points: Throughout the exercise, over half of all participants indicated that they had **no interest** in pursuing an offensive BW capability, with most of these participants still expressing an interest in biodefense. However, a sizeable number (>20) indicated that they had **slight interest** in pursuing an offensive capability. Only a small number of participants indicated **moderate or strong interest** in pursuing an offensive BW capability, but this number **increased** after participants were exposed to the first experimental inject (Russia's allegations). Following the response from the US and Ukraine, the number of participants expressing moderate or strong interest in pursuing an offensive capability **decreased**, but only slightly. These findings are displayed in Figure 1 below.

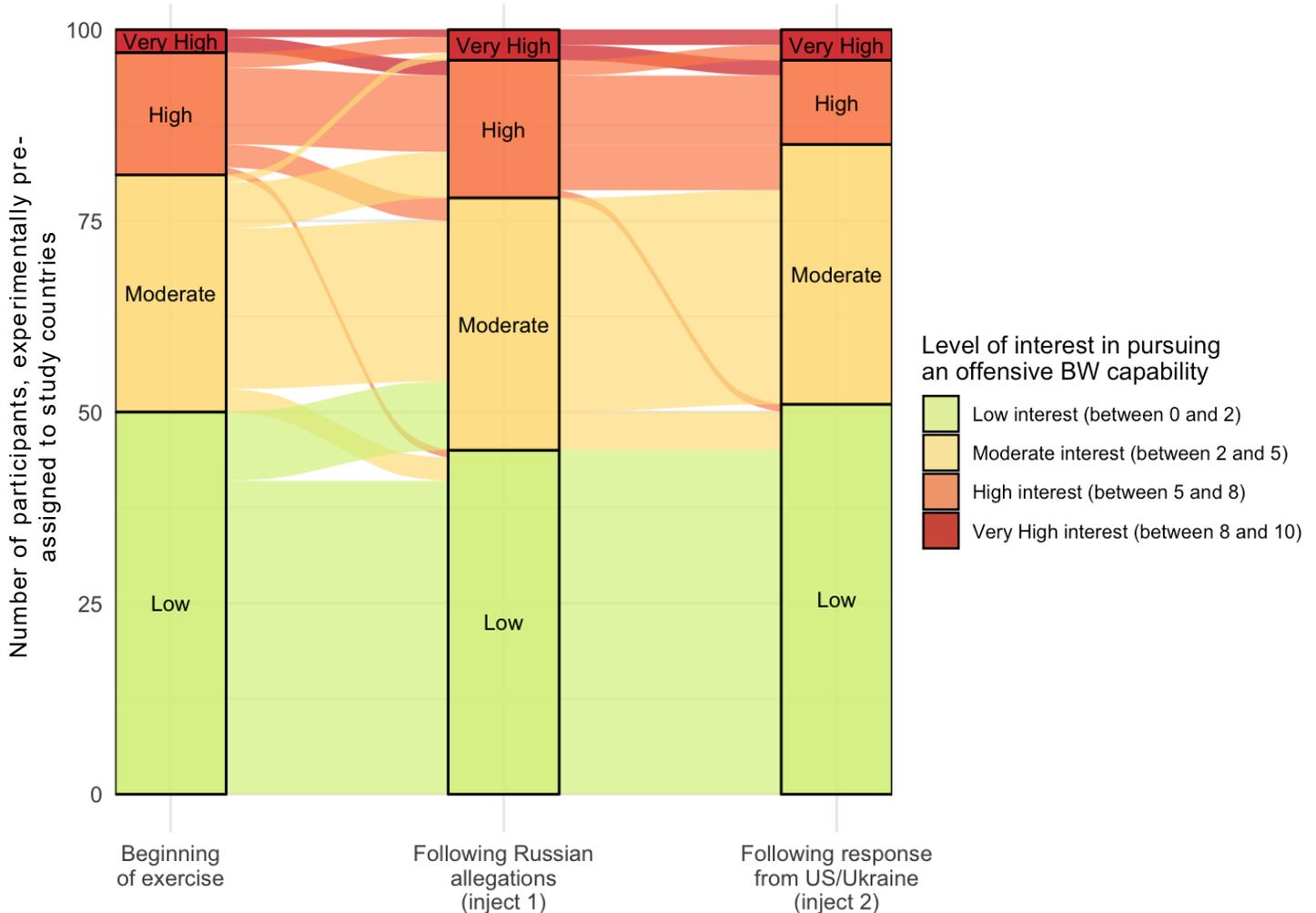
Figure 1: Participants' interest in pursuing activities related to BW at each of the three measured time points



Change in participants' interest in pursuing an offensive BW capability throughout the exercise: To better understand how interest in pursuing an offensive BW capability varied throughout the exercise, we grouped participants according to their ratings on the numeric (0 - 10) scale. We then plotted the number of participants in each group at each of the three measured time points (at the beginning of the exercise, following inject 1, and following inject 2). These groupings—and any movement between them—are displayed in Figure 2. As can be seen in the figure:

- For many participants, their interest in pursuing an offensive BW capability **remained at the same level** throughout the exercise.
- However, for some participants, their interest in pursuing an offensive BW capability **increased following the first inject**, with some of those expressing low or moderate interest at the beginning of the exercise expressing moderate, high, or even very high interest after exposure to Russia's allegations.
- For other participants, particularly those expressing high or very high interest at the beginning of the exercise, their interest in pursuing an offensive BW capability **decreased following the first inject**.
- Following the second inject, interest in pursuing an offensive BW capability decreased for many participants, though for some, **these decreases were small in scale**.

Figure 2: Change in participants' interest in pursuing an offensive BW capability throughout the exercise, with participants categorized according to their ratings on the numeric (0 – 10) scale.



Participants' explanations for their ratings: Participants were asked to provide short written statements explaining their level of interest in pursuing an offensive BW capability after each inject. Some of the more common explanations are described in the table below, alongside representative quotes.

Time point	Change in interest	Reason(s)	Representative quote
Following inject 1 (Russia's allegations)	Increased interest	Information in the inject raised concerns about a potential BW threat from the US and/or Ukraine	<i>"Our friends in Russia have shown considerable evidence that Ukraine and the US are engaging in offensive work... It's clear that we must now at least consider, seriously, how best to protect ourselves..."</i>
	Decreased interest	Perceived political costs of developing an offensive BW capability following Russia's allegations	<i>"Noting the tension between Russia and the West, my inclination is to maintain relationships with both—including on defense—and that it isn't wise to engage in risky biological research that stands to disrupt that relationship."</i>
	No change	Information provided in the inject did not change political priorities or commitment to nonproliferation norms and laws	<i>"[Our country] continues to abide by the rule of law and international norms, as set by the BWC. ... Although I take these claims seriously, it is crucial that we do not allow accusations [to] justify the abandonment of our principles."</i>
Following inject 2 (response from the US and Ukraine)	Increased interest	Perception that U.S. and Ukrainian activities still have dual-use potential; perceived ability to mask an offensive program as a defensive one	<i>"The US and Ukraine's activities, even if framed as peaceful, reveal the dual-use potential of biological research under defense programs, which I view as a strategic threat. Their reliance on public health narratives provides a model for plausible deniability, giving me cover to advance similar efforts under a defensive pretext."</i>
	Decreased interest	U.S./Ukrainian response perceived to be legitimate and trustworthy, alleviating concerns	<i>"The research conducted in Ukraine does seem targeted at reducing risks rather than developing weapons. It seems likely that the Americans are telling the truth..."</i>
	No change	Distrust of the US and its intentions	<i>"US evidence, in my view, simply added detail ... to the well-grounded accusations that Russian authorities made. American diplomats are skilled at manipulating international institutions like the UN to distort and mask [their] disdain of international norms and the lives of people who dare to live outside the imperialist orbit."</i>

Disinformation has the potential to alter state-level decision making about biological weapons

- Some participants' interest in pursuing an offensive BW capability increased following exposure to Russia's allegations, suggesting that for some state actors, **disinformation may increase the perceived utility or desirability of biological weapons.**
- These findings suggest that **disinformation can have tangible—and dangerous—impacts on global nonproliferation**, underscoring the need to invest resources in counter-disinformation efforts.
- In their written explanations for their ratings, participants often discussed—implicitly or explicitly—the geopolitical orientation of their assigned country, including their assigned country's relationship to the United States, Ukraine, and Russia. Such findings suggest that **the effects of disinformation may hinge, at least to some extent, on geopolitics.**

Debunking can mitigate some of the consequences of disinformation but should not be the only approach taken

- Many participants expressed a reduced interest in pursuing an offensive BW capability following the second inject (the response from the US and Ukraine). These findings suggest that **responding to and refuting false allegations (i.e., debunking) can mitigate some of the negative consequences of disinformation campaigns.**
- However, for some participants, their interest in pursuing an offensive capability did not decrease back to its starting level (i.e., to what it was at the beginning of the exercise), suggesting that **debunking may not always be 100% effective.**
- Many of these participants were assigned to countries with close ties to Russia and/or negative or strained relations with the United States, suggesting that **countries may respond to both disinformation and efforts to counter it along geopolitical lines.** Such responses can be described as a form of motivated reasoning, or reasoning that is driven by the desire to maintain pre-existing social or political beliefs and identities.²⁻³
- These findings highlight that while debunking can be an effective strategy, **other approaches to countering disinformation (including more pre-emptive strategies such as prebunking) should also be considered to ensure maximum efficacy.**

- ✓ **It is imperative that countries respond to false allegations fully and promptly**, particularly when such allegations rise to the level of a BWC Formal Consultative Meeting. Findings from this study suggest that debunking can be an effective way to respond to false allegations, albeit one that may not be 100% effective.
- ✓ **Debunking must be done carefully** and in a way that does not prompt countries to reject accurate information simply because it challenges pre-existing beliefs or geopolitical orientations. While this study found evidence that the response provided by the US and Ukraine was effective in some cases, it also appeared to trigger motivated reasoning in other cases.
- ✓ **Message tailoring techniques**, including aligning messages to audiences' core values, may help countries avoid triggering motivated reasoning.⁴ Aligning messages to the core values outlined in the BWC, including trust, transparency, and the importance of facilitating access to materials and technology for the peaceful use of biology, may be a particularly effective strategy.
- ✓ In addition to debunking, **other strategies to counter disinformation**, including pre-emptive refutation (i.e., prebunking) should also be utilized.

1. Ackerman GA, Behlendorf B, Peterson H, LaTourette J, Wetzel A, Clifford D. Investigating the Potential Strategic Implications of COVID-19 for Biological Weapons Pursuit: A New Expert Simulation. *Health Secur.* 2024;22(1):45-57. doi:10.1089/hs.2023.0029
2. Kunda Z. The case for motivated reasoning. *Psychol Bull.* 1990;108(3):480-498. doi:10.1037/0033-2909.108.3.480
3. Kahan DM. Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making.* 2013;8(4):407-424. doi:10.1017/S1930297500005271
4. Carpenter CJ. Cognitive Dissonance, Ego-Involvement, and Motivated Reasoning. *Annals of the International Communication Association.* 2019;43(1):1-23. doi:10.1080/23808985.2018.1564881