Summary

A nuclear exchange may have the potential to kill millions or billions of people, and possibly lead to human extinction.

In this post, I rank plausible nuclear exchange scenarios in terms of their potential to cause harm based on three factors: 1) The size of the involved countries’ nuclear arsenals; 2) The size of the involved countries’ populations; 3) The probability of the given nuclear exchange scenario.

Based on my rough prioritization, I expect the following nuclear exchange scenarios have the highest potential for harm:

1. Russia and the US
2. India and Pakistan
3. China and either the United States, India, or Russia
Project Overview

This is the first post in Rethink Priorities’ series on nuclear risks. In this post, I look into which plausible nuclear exchange scenarios should worry us most, ranking them based on their potential to cause harm. In the second post, I explore the make-up and survivability of the US and Russian nuclear arsenals. In the third post, I estimate the number of people that would die as a direct result of a nuclear exchange between NATO states and Russia. In the fourth post, I estimate the severity of the nuclear famine we might expect to result from a NATO-Russia nuclear war. In the fifth post, I get a rough sense of the probability of nuclear war by looking at historical evidence, the views of experts, and predictions made by forecasters. Future work will explore scenarios for India and Pakistan, scenarios for China, the contradictory research around nuclear winter, the impact of several nuclear arms control treaties, and the case for and against funding particular organizations working on reducing nuclear risks.

Toward a better understanding of nuclear risks

A nuclear exchange may have the potential to kill millions or billions of people, and possibly lead to human extinction. There have been many cases where nuclear weapons have almost been launched by mistake (Baum, de Neufville & Barrett, 2018).1 And if a nuclear exchange — started by accident or on purpose — were to escalate to a full-scale nuclear war, the nuclear detonations could lead to a nuclear winter, a state where soot launched into the atmosphere blocks out enough sunlight to cause a famine so severe and long-lasting that almost everyone on Earth could starve to death before its end.

Because there seems to be a non-negligible probability of a large-scale nuclear exchange, and because the stakes would be so high in the event that a nuclear exchange did escalate, many effective altruists believe reducing nuclear risks should be among the top priorities for the Effective Altruism (EA) community. For example, 80,000 Hours published a problem profile on nuclear security, giving a score of 15 out of 16 on Scale (though it scores relatively low on Solvability and Neglectedness).2

But my sense is that some details of the nuclear risks problem area aren’t well-understood by most EAs — for example, how bad nuclear war would actually be, the mechanisms behind nuclear winter, and where EAs that prioritize reducing nuclear risks should donate. In a number of upcoming posts, I’ll try to understand, in somewhat concrete terms, how much harm nuclear war would cause and how plausible nuclear risks are. One of the things I’ll do to better understand the risks posed by nuclear winter is review the implications of

---

1 See pages 27-32 for an itemized list of the US-Russia near misses (Baum, de Neufville & Barrett, 2018).
2 Note, they now recommend people use their in-depth interview with Daniel Ellsberg as a source of information on nuclear security.
recent academic literature that is interpreted by some as casting doubt on the science behind the nuclear winter phenomenon. Finally, I’ll also evaluate some of the work being done to reduce nuclear risks. In particular, I’ll focus on a recent treaty that’s been adopted by the United Nations, the Treaty on the Prohibition of Nuclear Weapons (TPNW), which would make the research and use of nuclear weapons illegal in all countries that ratify the treaty.

Focusing on the most troubling nuclear risks

In the first few posts, I’ll consider the probability and severity of several nuclear war scenarios, looking separately at the amount of harm that would be caused from both the short-termist perspective and the long-termist one.

It would be intractable to estimate the impacts of every imaginable nuclear war scenario. So I instead focus on nuclear war scenarios which I expect make up the majority of the expected harm that would be caused by nuclear war. These are scenarios in which:

1. *The specified countries have relatively large nuclear arsenals.* This is an important consideration for two reasons: first, because the size of the nuclear arsenal is a major factor in how severe the direct effects of a nuclear exchange are — more nuclear weapons can cause many more deaths. Second, whether a nuclear exchange leads to a nuclear winter depends in large part on the number and size of nuclear weapons detonated. It would take a lot of nuclear weapons to produce a nuclear winter severe enough to cause a worldwide famine that could lead to human extinction. Given that I’m most worried about nuclear winter scenarios that pose an extinction risk, I believe we should focus on nuclear exchange scenarios that would involve large nuclear arsenals.

2. *The countries involved have large populations.* This is an important consideration for similar reasons: first, because the population size of a country is an important factor in determining how many people could die as a direct result of the nuclear detonations. Second, whether a given nuclear exchange would lead to a nuclear winter depends on how much smoke is produced from the burning of cities during the exchange. Countries with larger, densely populated cities have much more flammable material. This means that a nuclear exchange involving densely populated countries would be more likely to lead to nuclear winter, all else equal.

3. *The specific conflict scenarios are reasonably probable.* This is important because the expected harm is higher in nuclear exchange scenarios that are more likely to actually happen, all else equal.

When I consider these factors for all 9 nuclear weapons possessor states, I get a rough ranking of various plausible nuclear exchange scenarios in terms of their expected dis-value:
The ranking is based on a crude scoring system that approximates the expected harm that would be caused by each scenario. Each of the factors that informs how terrible a given nuclear exchange would be — the size of the involved countries' nuclear arsenals, the size of the populations of the involved countries, and the probability of the specific scenario — were assigned a score of 1 (shaded in green), 2 (yellow), or 3 (red), where a factor with a score of 1 should worry us less, and a factor with a score of 3 should worry us a lot. For each scenario, the factor scores were summed together to produce the Expected Harm score.

Note that some endnotes are embedded in the table image but can be seen here: source for arsenal size;\(^3\) source for median war probability;\(^4\) note on non-state actors.\(^5\)

---

\(^3\) Arsenal data from the Bulletin of the Atomic Scientists' Nuclear Notebooks.
\(^4\) I believe the Project for the Study of the 21st Century (PS21) Great Power Conflict Report has several typos (Apps, 2015). I present what I believe to be the correct values (and the values I use in my analysis) here.
\(^5\) While a nuclear detonation by a non-state actor (terrorist) looks plausibly quite harmful in expectation, it'd be very difficult to analyze, as there's no single terrorism scenario to consider. I therefore leave a discussion of the potential harm caused by nuclear terrorism for future work.
Next Steps

Based on this rough prioritization, I’ll spend several posts looking at the amount of harm we would expect to see caused by the following nuclear exchange scenarios:

1. A scenario where the US and Russia use nuclear weapons, intentionally or accidentally, in a conventional nuclear exchange.
2. A scenario where India and Pakistan use nuclear weapons, intentionally or accidentally, in a conventional nuclear exchange.
3. A scenario where China uses nuclear weapons, intentionally or accidentally, in a conventional nuclear exchange with either the United States, India, or Russia.

Edits and Corrections

July 17 2019 — I replaced a simplified version of the table ranking potential conflict scenarios on the basis of nuclear arsenal size, population size, and scenario probability with a more detailed version. I also added a brief explanation of how the Expected Harm scores were calculated.

Credits

This essay is a project of Rethink Priorities. It was written by Luisa Rodriguez. Thanks to Peter Hurford, Marinella Capriati, Ida Sprengers, Marcus A. Davis, and Neil Dullaghan for their valuable comments. Thanks also to Matt Gentzel, Seth Baum, and Carl Shulman for providing guidance and feedback on the larger project. If you like our work, please consider subscribing to our newsletter. You can see all our work to date here.
Bibliography


Daniel Ellsberg on the creation of nuclear doomsday machines, the institutional insanity that maintains them, and a practical plan for dismantling them [Audio blog interview]. (2018, September 24). Retrieved April 14, 2019, from https://80000hours.org/podcast/episodes/daniel-ellsberg-doomsday-machines/


