Nuclear Challenges 2020 Evaluation and Learning Synthesis

ANNEX 1:
Summary of Methods

Table of Contents

Measurement Approaches: Strengths and Limitations

Overview

The Nuclear Challenges evaluation and learning framework was presented to John D. and Catherine T. MacArthur Foundation leaders in March 2019, coinciding with presentation of the Nuclear Challenges theory of change (TOC). The 2019 Evaluation and Learning Framework laid out an initial plan for data collection. As ORS Impact's data collection efforts took place between May 2019 and May 2020, methods were further developed and specified. This annex provides full details about the methods undertaken.

The Nuclear Challenges strategy is wide-ranging and—as outlined in the evaluation and learning framework—the measurement approach involved multiple data collection methods so as to sufficiently assess progress, capture evidence, and boost rigor. Along with review of strategy documents and grantee reports, data collection methods included:

- Key informant interviews
- Review of secondary sources, e.g., literature and media articles
- Congressional narrative analytics
- Congressional scorecard
- Grantee surveys
- Grantee focus groups

Data collection was designed to explore various aspects of the status of Nuclear Challenges' module- and strategy-level outcomes, including the breadth and depth of change; contributors to change, including critical landscape factors with potential to affect the likelihood, pace, and amount of change; and the status of impact measures. The multi-methods approach allowed for the description and analysis of progress toward identified outcomes and investigation of landscape factors, including those specifically identified in the evaluation and learning framework, and those that were emergent.

In addition to boosting rigor, the use of multiple methods also allowed for maximal utility of the data. As noted in the evaluation and learning framework, many outcomes identified in the TOC are qualitative rather than quantitative and therefore not easily observed or counted. To assess qualitative outcomes, ORS collected a broad set of data and triangulated as much as possible, both within a single method (e.g., comparing perspectives of key informants with different roles—Foundation staff and grantees, field leaders and experts, representatives of government—regarding particular scenarios or outcomes), as well as between different data sources (e.g., comparing findings from interviews, congressional narrative analysis, and the congressional scorecard). In addition to triangulation, we employed analytic techniques to boost rigor and add to the weight of evidence.

The Nuclear Challenges strategy is complex, and external conditions are dynamic and evolving with the potential to affect the strategy's implementation and progress. ORS sought to account for this complexity and the potential effects of external conditions by engaging in regular, close communications with Foundation staff; tracking and reviewing grantees' and experts' analysis of key events or situations; participating in meetings and work sessions with Foundation staff, the strategy's external advisors, and grantees; and independently tracking key events.

Nuclear Challenges investments toward the TOC's two core hypotheses (related to the regime and the fuel cycle) were focused through the strategy's five modules and associated components, and data collection activities also focused within these modules. Module-level findings were collectively analyzed to assess strategy-level progress against the two core hypotheses. As part of module-level analysis, we attempted to document and note progress on external conditions, landscape factors, and relevance to the strategy.

There were some challenges to data collection that ORS sought to manage during the data collection period:

- Gaining access to governmental representatives as key informants. To address this challenge, ORS primarily spoke with officials specifically named or referred by grantees and kept the interview protocol for governmental officials short and the time request to a minimum. Within this group of interviewees, it was especially challenging to access government officials above a senior level whose motivations are often important to determining decisions and actions. Grantees primarily referred ORS to U.S. government officials, although we were also able to interview a small number of European government representatives. We were unable to interview any government representatives in non-Western contexts, so findings may have a Western bias.
- Some key informants have sensitivities regarding candid sharing due to the field's proximity to national security. While the evaluation and learning framework focused on areas that can, in theory, be measured, sometimes the need for grantees to preserve confidentiality in order to protect their relationships and avoid disruption of ongoing efforts prevented them from candidly sharing certain information.
- Data collection focused primarily on short-term and intermediate module outcomes, and short-term strategy outcomes. This was reflective of the TOC in that a key intermediate-term outcome was not expected until at least 2025, and the long-term outcome following that. While there was robust data collection toward short-term and intermediate outcomes for most modules, some module and strategy components were not fully developed within the strategy at the time of data collection, for example, U.S. Weapons Policy and strategy-level capacity and effectiveness outcomes. For these components, there was less data collection and thus more limited analysis included in the synthesis report.

Data Collection Methods

Key Informant Interviews

Overview

The purpose of our interviews with knowledgeable key informants was to explore the status of and changes over time in Nuclear Challenges module- and strategy-level outcomes. Through semi-structured interviews with a broad range of key informants, ORS explored and documented the status of outcomes, relevant landscape factors, and the nature of change in outcomes and landscape factors over time. Speaking with those who hold different roles in the field allowed for a thorough understanding of outcomes and landscape factors, as well as the opportunity for triangulation across multiple perspectives, and comparison of interview findings with those arising from other methods. ORS completed 146 key informant interviews between May 2019 and May 2020.

ORS systematically analyzed the interviews using standard qualitative analytical practices to identify themes. Thematic analysis was intended to surface information about the status of and/or progress on identified module and strategy outcomes, relevant landscape factors (including hypothesized landscape factors and unanticipated or emergent factors), and the Foundation's influence and contributions.

Sampling

To ensure technical quality and alignment with the strategy and inquiry priorities, all of ORS' data collection instruments were reviewed by Foundation staff, including those from the Office of Evaluation and the Nuclear Challenges team. ORS also received input from Foundation staff regarding key informant samples for module and strategy inquiry.

The four main groups of key informants included:

- Nuclear Challenges **grantees**
- Nuclear Security Working Group (NSWG) **fellows** who are working or have previously worked in the office of a member of Congress (MOC)[1]
- Current or former **governmental officials**—i.e., MOCs, congressional staff, executive branch officials, military officials, or foreign officials—with whom grantees have interacted and/or whose positions afford them an informed perspective regarding nuclear issues

[1] With support from the MacArthur Foundation, George Washington University runs the Nuclear Security Working Group (NSWG), a program that promotes bipartisan discourse on critical nuclear security issues. The NSWG coordinates a congressional Fellowship that provides highly qualified up-and-coming national security professionals with an opportunity to spend a year working in an MOC's office.

- Non-grantee **experts** and others whose positions afford them an informed perspective regarding nuclear issues and the Foundation's influence and contributions regarding nuclear security

ORS customized interview guide for each group of key informants, tailored to the specific input that group was best positioned to provide. (See separate Annex that includes all data collection instruments.)

Strengths and Limitations

Table 1 lists sampling approaches used to identify grantees, fellows, officials, and experts who were invited to participate in interviews, along with their strengths and limitations.

Table 1. Key Informant Sampling Approaches, Strengths, and Limitations

Group	Sampling Approach	Strengths & Limitations
Grantees	• With information from the Foundation's grants management services and the NC program team, ORS proposed a sample of interviewees for each module based on how grantees' work aligned with a given module. • For grantees whose work aligned with multiple modules, the NC team helped determine which module(s) grantees were best positioned to speak to.	*Strengths* Grantees are knowledgeable experts and typically willing to be interviewed. Grantees are well positioned to provide insights about: (a) the landscape surrounding the strategy, (b) strategy and module outcomes, and (c) the Foundation's unique influence and contributions. In addition, grantees reflect a large proportion of the civil society organizations working toward nuclear risk reduction. *Limitations* Desires to maintain Foundation support could motivate grantees to overstate the influence, effectiveness, and/or significance of their efforts. Additionally, the number of grantees associated with each module varies; the potential sample for some modules was quite small, for example, the total number of grantees within the U.S.-Russia module is <5. Small samples limit the ability to triangulate multiple perspectives.
NSWG Fellows	• ORS identified individuals who served as fellows in the office of a MOC between 2017 and 2020.	*Strengths* Like grantees, fellows are well positioned to provide insights about the landscape surrounding the strategy, and outcomes for certain modules, especially the U.S. Congress module. *Limitations* A desire to maintain Foundation support for the fellowship program could lead fellows to overstate the influence, effectiveness, and/or significance of their efforts. ORS was not able to speak directly with the MOCs in whose offices fellows served nor their staff, and thus there is limited ability to confirm agreement between fellows' and officials' perspectives.

Group	Sampling Approach	Strengths & Limitations
Officials	• To identify potential key informants, ORS implemented a snowball sampling approach, wherein some interviewed grantees were asked to identify officials with whom they had had direct engagement, or who they knew to be knowledgeable about grantees' efforts, module and strategy outcomes, and landscape factors, and were willing to participate in an interview. • ORS (or grantees) contacted identified officials to pursue interviews using convenience sampling.	*Strengths* Officials can provide unique and useful insights about the landscape surrounding the strategy and the status of outcomes, as well as objective input about the influence and contributions of grantees and the Foundation toward change in outcomes of interest. *Limitations* The total number of officials with whom grantees regularly interacted is small, and because grantees were given freedom to opt out of suggesting officials if it would in any way jeopardize their current relationships or disrupt or otherwise negatively impact ongoing work, the overall sample was small. Outreach to officials was halted from March–April 2020 amidst concerns that officials would be addressing more immediate issues related to the coronavirus pandemic.
Experts (Non-grantees)	• With input from the NC team and grantees, ORS identified approximately 60 non-grantee experts, including program officers representing other nuclear funders, practitioners (those who hold positions in non-governmental or private sector entities), and researchers (those affiliated with academic institutions or think tanks). Some experts are former MacArthur Foundation grantees. • ORS reviewed proposed key informants with the NC team before finalizing the sample to include 41 of the 60 experts initially identified.	*Strengths* Experts are well positioned to provide insights about the landscape surrounding the strategy and the status of outcomes. They may also provide alternate perspectives from grantees regarding the degree to which the Foundation's efforts have/have not influenced or contributed toward outcomes of interest. *Limitations* The extent of experts' knowledge and ability to speak to the landscape, status of module or strategy outcomes, and/or the Foundation's influence and contributions is not known in advance; these factors could affect the relevance and quality of data gathered through interviews. Experts have no obligation to participate in interviews, which could affect response rates. Outreach to external experts was halted during March–April 2020 amidst concerns about their capacity amidst the coronavirus pandemic.

Timeframe

Key informant interviews began in May 2019 and were paused in May 2020 as the ORS team turned attention to analysis and synthesis.[2] ORS completed 146 interviews as follows:

- Of 118 individual grantees, 85 were contacted and 77 grantee representatives participated in 87 interviews as 10 grantees completed two interviews for two different modules (90 percent response rate)[3][4]
- 16 of 18 fellows were contacted; of those 16, 14 were interviewed (89 percent response rate)
- 21 of 40 identified U.S. and European government officials were interviewed (52 percent response rate)
- 24 of 41 identified external non-grantee experts were interviewed (58 percent rate)

[2] Eight key informant interviews were undertaken in June 2018 before the development of the evaluation and learning framework. As those interviews were not related to the evaluation of the strategy, the data was not analyzed for the purposes of the synthesis report and the interviews are not included in this count.

[3] In some cases, grantees were intentionally left out of the sample per the NC team's preference, e.g., because a grant had recently or was about to come to an end, or for reasons like medical leave. In one case (the U.S.-Russia module), two grantees were excluded from the sample because the Foundation's institutional policy prohibits contact with Russia-based organizations.

[4] Of grantees identified and approached as potential key informants, eight were not interviewed. Of those, five did not respond to ORS' request, and three agreed to participate but ORS was unable to complete the interviews. ORS' practice is to contact potential key informants with an explanation of purpose and a request for interview participation, and to recontact up to three times if the individual does not respond to the first request.

Secondary Sources

To understand outcomes, impacts, and the landscape within and across modules and the strategy, and to document and track change in specific areas, ORS has continually identified relevant secondary sources—reports, briefs, white and gray literature, and media articles—by:

- Conducting targeted digital searches to identify relevant sources on topics of interest, such as searches of the organizational websites of key informants, searches for output authored or produced by grantees or non-grantee experts, and output that addresses key topics associated with module or strategy outcomes
- Asking the Nuclear Challenges program team and key informants to recommend or share relevant reports, briefs, articles, or other secondary materials that they authored and/or that address topics of interest

ORS systematically reviewed secondary source documents using standard qualitative analytical practices to identify themes. Thematic analysis was intended to surface information about the status of and/or progress on identified module and strategy outcomes, relevant landscape factors (including hypothesized landscape factors and unanticipated or emergent factors), and the Foundation's influence and contributions. (See separate Annex for a full list of secondary sources reviewed.)

Congressional Narrative Analytics

Overview

Narratives represent the distillation of various expressed arguments and opinions within discourse into a cohesive story or worldview. Narrative analytics is a systematic approach to better understanding these narratives by applying a combination of social scientific approaches and data science. Narrative analytics can map, track, and measure discourse to surface certain narratives, learn who expresses particular narratives and with what frequency, and see how narratives or their expressions shift over time.

In the congressional landscape, narratives are evident in MOCs' individual expressions of their opinions and arguments. To understand these narratives, ORS implemented congressional narrative analytics in partnership with a subcontractor, Protagonist.[5] With input from ORS and the NC team, Protagonist undertook a two-phased inquiry. **Phase 1** mapped the narratives evident in MOCs' discourse output, identified the proportion of discourse that addressed nuclear topics compared with other topics, and surfaced changes in narrative over time. **Phase 2** involved two analytic deep dives: an analysis of key outcomes/measures identified in the evaluation and learning framework as they relate to MOCs, and an analysis of the identified top solutions and top countries discussed by policymakers from 2016–2019.

Given that the majority of MOCs are active on Twitter, narrative analytics relied on Twitter as a mechanism for collecting data such as public statements and relevant links, traditional media references, congressional speeches, and other published materials.

Strengths and Limitations

A strength of this approach is the ability to generate a deep and expansive picture of narratives in a given landscape. A limitation lies in the extent to which the approach can gather accurate and complete data. For example, since not all nuclear narrative is public, it is possible that Protagonist's narrative analytics, which relied on review of public statements (via Twitter posts), may not capture the full and total narrative regarding concepts of interest. In addition, Twitter behavior is selective and potentially biased, in that the user is choosing to Tweet with an eye toward reflecting their public position, similar way to how a user might think about media engagement. Thus, Twitter posts may not offer a perfect reflection of the user's engagement on a given issue.

Timeframe

The congressional narrative analytics dataset reflects data from January 1, 2016 to November 30, 2019. While it is conceivably possible to gather Twitter data back to March 2013, MOCs' participation on Twitter

[5] Protagonist is currently assisting the MacArthur Foundation's Climate Solutions team to explore and document how political discourse on climate change is evolving, and whether and how the Foundation's strategy is contributing to documented changes.

was inconsistent prior to 2016 and thus there are limited opportunities for comparative analyses, such as assessing growth of individual MOC's public sentiment and participation regarding nuclear issues.

Congressional Narrative Analytics - Phase 1

Protagonist's initial analyses focused on answering two questions:

- To what extent does recent **congressional public output focus on nuclear issues** compared to other policy topics, especially among specific MOCs of interest?
- Within the MOCs' public output that is focused on nuclear issues, what are **notable narratives being espoused**, and to what extent is there awareness of and/or support for specific solutions, particularly those that are aligned with NC outcomes (e.g., solutions that reflect commitment to nonproliferation, disarmament, and peaceful use)?

The first question was answered via a **topical analysis**. This work involved the following steps:

- Protagonist developed a custom dataset focused on MOC public output from 2016–present which captures Twitter output (including links to statements and speeches) from official MOC and congressional committee handles. Drawing on over 700 of these Twitter accounts from 2016 through 2019,[6] Protagonist analyzed over 1.7 million data points. Data were filtered using machine learning and natural language processing software, yielding a dataset of about 1 million rows.
- From this dataset, Protagonist identified and measured the prevalence of key policy topics (e.g., climate change, healthcare, etc.) present in MOC public output, and assessed how these have changed since 2016.
- Protagonist analyzed a subset of the data capturing public output from specific MOCs and compared it to the entire MOC dataset to determine if specific MOCs are more/less focused on nuclear issues and if this focus is growing.[7]

To answer the second question, Protagonist sought to establish the **narrative baseline** by analyzing the portion of the dataset focused on nuclear issues to identify notable narratives. This work involved the following steps:

- To map the landscape of MOC narratives about nuclear issues, Protagonist implemented a query of nuclear terms (identified with input from ORS and the NC team) to identify about 6,500 posts that were both nuclear-related and rich in narrative content. (See the full list of terms queried in Annex 1A.)

[6] A Twitter post was only included in the dataset if its posting date was within the date range where the MOC was in office.

[7] The subset of specific MOCs was identified with input from the NC team; this subset includes those MOCs who have hosted or expressed interest in hosting an NSWG Fellow, those who serve on key committees where nuclear issues are discussed (e.g., the House Armed Forces Committee, the Senate Foreign Relations Committee), and members of the Senate who are running for President.

- From this portion of the dataset, Protagonist analyzed MOCs' awareness of, engagement with, and support for nuclear issues, with particular attention to the three regime principles of nonproliferation, disarmament, and peaceful use of nuclear material—a focus of the NC strategy.
- Protagonist then measured whether the narrative reflects policy solutions and, if so, the level of MOCs' awareness of, support for, or opposition to solutions and changes in awareness and support over time. (Policy solutions were identified with input from ORS and the NC team; examples include the reauthorization of New START and the closure of the Savannah River MOX facility.)

Narrative Analytics - Phase 2

Based on Phase 1 analysis, Protagonist undertook two deep-dive analyses: an analysis of key outcomes/measures identified in the evaluation and learning framework as these relate to MOCs; and an analysis of the identified top solutions and top countries discussed by policymakers from 2016–2019. The **analysis of key outcomes and measures** included:

- Assessment of the narrative volume, consistency, and alignment for MOCs on four key committees: House Armed Services Committee; Senate Foreign Relations Committee; House Foreign Affairs Committee; and House Foreign Affairs Subcommittee on Asia, the Pacific, and Nonproliferation. These committees have the highest volume of nuclear posts (as represented by MOCs on these committees) in addition to the House Appropriations Committee.
- Assessment of the narrative volume, consistency, and alignment for MOCs who have hosted a NSWG fellow.
- Identification and assessment of MOCs who champion dialogue on nuclear issues in relation to key countries (e.g., U.S.-Russia); MOCs who support open lines of communication, a peaceful and diplomatic approach, etc. (associated with U.S.-Russia module outcomes); and/or MOCs who demonstrate increased awareness of analysis and recommendations related to solutions for Iran and North Korea (associated with Tough Cases module outcomes).
- Identification and assessment of MOCs who speak most often about nuclear issues, including their common and distinguishing characteristics; their alignment with different narratives including those aligned with regime principles, as well as to end use and eliminate stockpiles; and their openness to any new framing of nuclear policy issues (associated with strategy and U.S. module outcomes).
  - The method for identifying and assessing MOCs who speak the most (in terms of volume and consistency) on nuclear issues, who were classified as either aligned or unaligned with a specific narrative, e.g., 'Upholding Regime Principles'.[8] Those classified as unaligned tended to

[8] Via analysis, the Protagonist team identified several unique narratives. These are detailed in a separate technical report, available on The Loop.

support different narratives, e.g., 'Maintain U.S. Strategic Position'. Based on Phase 1 data, those MOCs were classified as at least one standard deviation above the mean in both volume and consistency of nuclear posts (as compared with the yearly average). In raw data terms, this means MOCs must have Tweeted at least 4.23 months on average about nuclear and 1.34 percent of their total discourse must have been dedicated to nuclear on average each year. Protagonist identified 32 MOCs who fit this category during the period 2016–2019.[9] The analysis included:

- Mapping the percentage of these 32 MOCs' narrative focus dedicated to specific narratives, including 'Upholding Regime Principles' and 'Maintain U.S. Strategic Position' (and related subcategories), 'Nuclear Energy' and 'Nuclear Waste and its Impacts,' and other nuclear narratives. This analysis also presents the total number of nuclear Tweets by narrative.
- A view of narrative over the period 2016–2019, as well as by year to highlight any changes over time by individual MOCs.
- Analysis of MOCs whose narratives diverge when it comes to 'Maintain U.S. Strategic Position' and 'Upholding Regime Principles' to see if there are identifiable themes and/or narratives.

The analysis of top solutions and countries discussed by policymakers included:

- Top solutions: The top five solutions based on Phase 1 findings were JCPOA, No First Use Bill, INF Treaty, IAEA, and New START.[10] Analysis of the top five solutions included volume trends, specific themes, policy champions and detractors, and related insights. (See examples of solutions listed in Annex 1B.)
- Top countries: This inquiry focused on the top six countries from the percentage of total posts where a country is mentioned: Iran (24 percent), North Korea (10 percent), Russia (4 percent), Saudi Arabia (1.6 percent), China (1.5 percent), and Israel (1.3 percent). The analysis allowed cross-referencing to the top solutions (e.g., JCPOA and Iran) as well as findings relevant to the outcomes and measures deep dive.[11]

[9] There was an option to increase the number of MOCs in the champions analysis, by changing from the 'one standard deviation above the mean' criteria that is currently in place. However, expanding the number of MOCs risked having a group of MOCs that is less indicative of champion status, and it was considered more valuable to maintain a strict boundary around the group.

[10] Other solutions (e.g., Treaty on Open Skies, the National Defense Authorization Act, Nuclear Posture Review, Nunn-Lugar Cooperative Threat Reduction, the Ban Treaty, NPT) represented very low volume, and were unlikely to yield any robust findings.

[11] Other specific countries (such South Korea, Japan, Germany, and India) were not examined through the Protagonist inquiry as the preliminary analysis found that the data was not sufficiently substantive to warrant a deep dive.

Congressional Scorecard

Overview

Congressional scorecards are tools used to rank sitting legislators (or candidates for legislative office) on their voting record. Scorecards are one way to assess how legislators support or champion certain types of policies. Typically aggregated on an annual basis, scorecards track and document the actions of individual legislators and identify the extent to which their actions are either supportive of or not supportive of a policy issue or set of issues. With scorecards, it is possible to examine trends in congressional engagement with or support for issue(s) over time. Because of the specialized nature of this approach, ORS subcontracted with Foreign Policy for America (FP4A) to develop a dataset that informs a custom scorecard that allows for assessment of the nature and extent of legislative action on nuclear issues, and changes in legislative action over time.

Strengths and Limitations

The strength of this approach is that scorecards provide a clear picture of legislative actions and how these are changing over time, especially when combined with findings from other approaches, such as narrative analytics and key informant interviews. A limitation is that the scorecard focuses only on co-sponsorship and votes; a wider range of legislative behaviors, including floor statements, letters, and other actions that indicate champion status are not included and thus the scorecard may not show the full picture of legislative actions.

Timeframe

FP4A completed their work in two phases. During the first phase, between November 2019 and January 2020, FP4A drew on its biennial Legislative Scorecard to develop a customized scorecard covering the 114th, 115th, and 116th Congresses (roughly aligned with the period 2016–2019). During the second phase, between March and June 2020, FP4A expanded their scorecard to cover the 111th, 112th, and 113th Congresses (roughly aligned with the period 2013–2016). The resulting dataset permitted FP4A's analysis of the scorecard data covering six Congresses.

Approach

FP4A's analysis of scorecard data resulted in findings that triangulate with ORS Impact's assessment of outcomes for the U.S. Congress module and the congressional narrative landscape.

FP4A's work to develop their congressional scorecard and analyze data had four steps:

- **Step 1:** Identify the universe of legislation related to nuclear issues for the 111th, 112th,113th, 114th, 115th, and 116th Congresses. Categorize each bill into 12 legislative categories (identified by FP4A) based on its content. The universe of legislation includes legislation that aligns with the goals of the NC modules, as well as legislation that relates to other topics that have not been a focus of the

strategy, including de-emphasized regime topics (identified in Appendix B of the March 2019 Evaluation and Learning Framework).

- Step 2: Code each piece of legislation as either "consistent" or "inconsistent" with the principles of disarmament, nonproliferation, and peaceful use, based on the guidance provided by ORS and information provided by the NC team. Legislation that is neither consistent nor inconsistent with the principles was coded as "complex."
- Step 3: Score each legislator based on the percentage of times their decisions aligned (or were consistent) with the principles. FP4A created two scores for each legislator: one that reflects the total percentage of aligned decisions, and one that reflects the percentage of aligned decisions on key legislation (i.e. legislation that received co-sponsorship by one tenth of the body). The 1/10 co-sponsorship threshold was selected to identify legislative action that had gained a significant level of traction within the whole Congress. Legislation coded as "complex" is not included in the scoring.
- Step 4: Conduct deeper analysis of the full, six-Congress scorecard dataset. FP4A assessed changes in legislation and aligned decisions over time, including whether/how the percentage of aligned decisions within Congress changed over time, and what has driven these changes, focused on the following:
  - Who changes: Among which MOCs are there changes in aligned decisions over time? What changes are notable among different groupings of legislators (e.g. moderate Democrats, narrative champions, etc.)?
  - What explains changes: What drives legislators' aligned decision making? Are there clear patterns in legislators' decision making—are there differences due to party, chamber, or other MOC characteristics that may be relevant? If there are patterns, what might these patterns indicate about the breadth, nature, and/or extent of congressional support for NC-aligned outcomes?
  - Trends in nuclear legislative categories over time, including whether/how there are changes in the nuclear-related legislative categories over time. FP4A assessed and provided a qualitative description of:
    - The overall amount of legislative action in specific categories that changed over time
    - The amount of aligned legislative action over time, i.e., the proportion of legislative action in each category that aligns with the goals of the NC strategy
    - The major issues addressed in each category and the legislative story for each category

The congressional scorecard was implemented primarily to assess outcomes within the U.S. module; however, findings are also relevant to assessment of outcomes in other modules, e.g., Iran, North Korea, and Possessor/Non-Possessor Tensions.

Grantee Surveys

Overview

ORS implemented surveys with selected groups of grantees at different points in time, which provided information that helped ORS to assess module and strategy outcomes and the landscape.

Surveys were implemented with those who participated in three Foundation-supported communications workshops in 2018 and 2019. ORS also supported the NC team in developing and implementing a survey in conjunction with two grantee convenings held in 2019 to gather feedback from grantees about the NC strategy.

In April 2020, ORS implemented a survey with all NC grantees regarding three areas of inquiry:

- Strategy-level capacity and effectiveness outcomes
- U.S. Weapons Policy (USWP) module outcomes
- Line of sight to all module and strategy long-term outcomes

Grantees' area of work determined the survey questions they were best positioned to answer. Depending on their area of work, grantees received one of four versions of the survey, shown in Table 2. Response rates for the all-grantee survey are shown by survey version in Table 2; Table 3 provides the response rates by the three areas of inquiry.

Table 2. Response Rates by Version of the Grantee Survey Received

Survey Version	Responses (n)	Response Rate
All survey questions–i.e., capacity/effectiveness outcomes, USWP outcomes, long-term outcomes	6 out of 20 (excludes 3 partial completions)	30%
Questions re: capacity/effectiveness outcomes and long-term outcomes	21 of 45 (excludes 1 partial completion)	47%
Questions re: USWP module outcomes and long-term outcomes	4 of 15 (excludes 1 partial completion)	27%
Questions re: long-term outcomes only	22 of 38	58%

Table 3. Response Rates by Grantee Survey's Area of Inquiry

Area of Inquiry	Responses (n)	Response Rate
Capacity/effectiveness outcomes	30 of 65	46%
USWP module outcomes	12 of 35	34%
Long-term outcomes	53 of 118	45%

Strengths and Limitations

This approach was effective for capturing grantees' assessment of the extent of progress toward outcomes related to capacity and effectiveness, U.S. Weapons Policy module, and long-term module outcomes. However, due to low response rates, the data offer limited usefulness for telling the story of progress and grantees' contributions. Limitations precluded more nuanced exploration of survey data.

Timeframe

ORS implemented surveys with grantees who participated in three Foundation-supported workshops in 2018 and 2019, as well as for two grantee convenings held in July and September 2019. ORS also implemented an all-grantee survey in April 2020.

Grantee Focus Groups

Overview

Focus groups are a qualitative method that is essentially a group interview. Focus groups intentionally bring together certain participants to participate in an open discussion where they are asked about their perceptions, opinions, and beliefs regarding a certain subject. Questions are asked in a manner that encourages interaction among group members.

Strengths and Limitations

Like interviews, one strength of focus groups is that they are a useful way to obtain individual and group perspectives. Focus groups also offer the opportunity to seek clarification or detail regarding certain questions or emerging findings. Since focus groups can be implemented when groups are already scheduled to come together, they can be time- and cost-effective compared to individual interviews.

Some limitations of this approach are the potential for focus groups to elicit irrelevant discussion or disagreements between participants, which can distract from the main focus, and the potential for one or a small number of participants to dominate discussion, crowding out other important perspectives. Finally, because focus group participants are often identified via convenience sampling or self-selection, they may not be fully representative of all viewpoints.

Timeframe

In the fall of 2019 and spring of 2020, ORS implemented a series of four focus groups with approximately 30 grantees whose work connects with the Iran or North Korea component of the Tough Cases module, or the Possessor/Non-Possessor Tensions module. Focus groups were scheduled in conjunction with planned grantee convenings. They presented an opportunity to gather additional data from grantees (who had previously participated in key informant interviews) regarding the status and change in outcomes and landscape factors, and more deeply explore past and emerging findings.

Analytical Approaches

Modules, Module Components, and Impact Measures – Early Analysis

To review and synthesize data from grantee interviews and early review of secondary sources, ORS engaged in standard qualitative analysis via thematic coding and analytic dialogue sessions, and then produced technical reports based on these analyses. These early findings reports were reviewed by the Office of Evaluation and the NC program team. Subsequent discussion allowed the NC team to digest and apply findings, and to provide feedback which enriched interpretation of findings.

Between May 2019 and May 2020, ORS produced the following technical reports:

- A report that addressed outcomes and landscape factors relevant to the fuel cycle, encompassing the following modules and module components: U.S. Fuel Cycle, Global Fuel Cycle, and Threshold Countries: Fuel Cycle.
- Two separate reports that each addressed outcomes and landscape factors relevant to the U.S. Congress and U.S.- Russia components of the U.S. module.
- Two separate reports that each addressed outcomes and landscape factors relevant to the Tough Cases module components: North Korea and Iran.
- A report addressing outcomes and landscape factors relevant to the Possessor/Non-Possessor Tensions module.
- A report characterizing the baseline period preceding the Nuclear Challenges strategy, 2009–2015.
- Four brief reports addressing particular aspects of the external nuclear landscape, including the philanthropic funding environment, emerging technology, public awareness and public opinion, and bilateral and multilateral relationships.
- A report that presented status and change regarding impact measures.

Data gathered in relation to other module components, including Threshold Countries: Weapons and United States Weapons Policy, were not summarized in standalone technical reports. As a relatively new area of investment, grantmaking within the Threshold Countries: Weapons module has been limited to date, with only three organizations receiving grant support specific to this area. However, analysis of data gathered from these grantees was included in the wider cross-strategy synthesis. Data specific to outcomes and landscape relevant to the United States Weapons Policy module component were primarily collected through the all-grantee survey, but due to a low response rate, data limitations prevented ORS from presenting a robust story of progress and grantees' contributions regarding module outcomes.

Approach to Synthesis

Synthesis efforts were undertaken during May–June 2020.

In order to analyze, triangulate, and synthesize the wide range of data collected, the ORS evaluation team held a series of dialogic sessions.[12] The first session incorporated analysis of data and findings not included in the standalone technical reports, such as those from:

- Interviews with officials
- Interviews with key informants
- Grantee survey
- Grantee focus groups

The next sessions focused on synthesis of data relevant to each of the two pathways within the TOC: the fuel cycle and the regime. Through synthesis, we sought to roll up findings from across modules and the strategy as follows:

Fuel Cycle

- Synthesis drew on interview findings from Global Fuel Cycle, U.S. Fuel Cycle, and Threshold Countries: Fuel Cycle modules and components, plus related data such as secondary sources, findings from interviews with key informants, and grantee survey findings.

Regime

- Examination of elements of U.S. leadership and engagement regarding nuclear issues drew on interview findings from the U.S. Congress module component, interviews with fellows, congressional narrative analytics and congressional scorecard data and related analyses, secondary sources, findings from interviews with key informants, grantee survey findings, and focus groups with grantees.
- Examination of bilateral and multilateral dialogue drew on interview findings from the Tough Cases module components (Iran and North Korea), Possessor/Non-Possessor Tensions module, Threshold Countries: Weapons module component, and related data such as narrative analytics, findings from interviews with key informants, secondary sources, grantee survey findings, and focus groups with grantees.

During ORS' analysis sessions, themes from the various data sources were identified, triangulated against each other, and weighted in relation to their quality. Based on the weight of evidence, the ORS team also sought to build consensus regarding indicative themes and conclusions.

[12] Dialogic sessions are a common qualitative analytical practice. ORS teams use these sessions as an opportunity for multiple individuals involved in (different aspects of) data collection to bring forward their unique perspectives, engage in discussion about findings, and build consensus regarding meaningful themes and weight of evidence.

Annex 1A. Congressional Narrative Analytics: Nuclear Terms Queried[13]

nuclear*, denuclear*, #nuclear, nonproliferation, non-proliferation, “nonproliferation”, “peaceful use”, uranium, “enriched uranium”, “highly-enriched uranium”, “HEU minimization”, “ HEU ”, “low-enriched uranium”, “ LEU ”, “uranium-235”, “uranium-233”, “uranium-238”, plutonium, “plutonium-239”, “plutonium-240”, “first-use”, #FirstUse, #NoFirstUse, #RestrictFirstUse, “first strike”, first-strike, “possessor state”, counterproliferation, “counter-proliferation”, “MOX facility”, “MOX fuel”, “MOX services”, “mixed oxide”, “mixed-oxide”, “mixed plutonium-uranium oxide”, “W-UM stockpiles”, “weapons-usable material”, “weapons usable material”, “weapons-usable materials”, “weapons usable materials”, “weapons-usable nuclear material”, “weapons-grade material”, “weapons grade material”, “fissile material”, “atomic bomb”, “atom bomb”, “a-bomb”, “thermonuclear”, “hydrogen bomb”, “h-bomb”, “fission bomb”, “ ICBM ”, “ ICBMs ”, “intercontinental ballistic missile”, “ SLBM ”, “submarine launched ballistic missile”, “reactor”, “ NPT ”, “treaty on the non-proliferation of nuclear weapons”, “non-proliferation treaty”, “NPT Review Conference”, “NPT framework”, “ TPNW ”, “Treaty on the Prohibition on Nuclear Weapons”, “Ban Treaty”, “ CTBT ”, “Comprehensive Nuclear Test Ban Treaty”, “Test Ban Treaty”, “ NWFZ ”, “Regional Nuclear Weapons Free Zone”, “WMD free zone”, “Weapons of Mass Destruction Free Zone”, “Treaty on Open Skies”, “Open Skies”, “INF Treaty”, “Intermediate-Range Nuclear Forces Treaty”, “New Strategic Arms Reduction Treaty”, “New START”, “Strategic Arms Reduction Talks”, “ JCPOA ”, “Joint Comprehensive Plan of Action”, “Iran Deal”, #IranDeal, “ CEND ”, “Creating an Environment for Nuclear Disarmament”, “ IAEA ”, “International Atomic Energy Agency”, “UN Conference on Disarmament”, “annual conference on disarmament”, “U.N. Conference on Disarmament”, “Congressional Nuclear Security Working Group”, “Nuclear Posture Review”, “UN First Committee”, “U.N. First Committee”

[13] Nuclear terms were identified with input from ORS and the Nuclear Challenges team. Based on these terms, Protagonist implemented a query and surfaced approximately 6,500 Twitter posts that were both nuclear-related and rich in narrative content.

Annex 1B. Congressional Narrative Analytics: Policy-Related Solutions Queried

A. Multilateral Treaties

- Treaty on the Non-Proliferation of Nuclear Weapons (NPT) or NPT Review Conference 2020 or NPT Framework
- Treaty on the Prohibition on Nuclear Weapons (TPNW) or Ban Treaty
- Comprehensive Nuclear Test Ban Treaty (CTBT)
- Regional Nuclear Weapons Free Zone (NWFZ) or Weapons of Mass Destruction (WMD)-free zone
- Treaty on Open Skies or Open Skies

B. U.S.-Russia Treaties

- New Strategic Arms Reduction Treaty or New START
- Intermediate-Range Nuclear Forces Treaty or INF

C. Iran/United Nations Security Council Resolution

- Joint Comprehensive Plan of Action (JCPOA)

D. Institutions

- International Atomic Energy Agency (IAEA)
- United Nations Conference on Disarmament

E. U.S.-Specific Initiatives/Policy/Legislation

- Creating an Environment for Nuclear Disarmament (CEND)
- No First Use Bill
- Nuclear Posture Review