



EVALUATING ADVOCACY
FOR POLICY IMPLEMENTATION:
A Pilot of Two Tools

LEARNING REPORT

CONTENTS

EXECUTIVE SUMMARY	1
I. INTRODUCTION	5
II. THE CONTEXT	5
III. THE GENDER SCORECARD FOR U.S. ADMINISTRATION OFFICIALS	8
IV. QUALITY OF DISCOURSE (QOD) CONTENT ANALYSIS	14
V. LESSONS AND RECOMMENDATIONS FOR FUTURE APPLICATIONS	22
VII. CONCLUSION	26
APPENDIX A: GENDER SCORECARD FOR ADMINISTRATION OFFICIALS	27
APPENDIX B: SUMMARY OF CONTENT ANALYSIS CODES	30

This report is the product of a sustained and fruitful collaboration between CARE USA and the Aspen Institute. Dr. Susanna Dilliplane, David Devlin-Foltz, Alex Gabriel, and other Aspen Institute colleagues worked with various staff, fellows, and interns at CARE USA over the past three years to develop and pilot the methodology and tools featured in the report. Dr. Dilliplane authored this report in collaboration with Milkah Kihunah, Dr. Nidal Karim, Elizabeth Brezovich, and other staff at CARE USA who provided invaluable comments and insights.



EXECUTIVE SUMMARY

From 2009 to 2016, CARE USA and the Aspen Planning and Evaluation Program (APEP) partnered to develop a set of tools for evaluating policy advocacy. Our aim was to track progress toward CARE’s advocacy goals, assess CARE’s influence, and provide insights to inform CARE’s strategy. Our efforts were part of the nascent, but growing, field of advocacy evaluation, which grapples with the challenge of measuring advocates’ contributions to complex policymaking and policy implementation processes. Policy implementation poses special evaluation challenges.

The purpose of this report is to share what we learned from our experience piloting two tools focused on CARE’s Gender and Empowerment advocacy in particular: a Gender Scorecard for U.S. Administration Officials, and a Quality of Discourse tool. The goal of these two tools was to help CARE monitor whether and how the US Government was implementing two key policies on gender-based violence (GBV) and child, early, and forced marriage (CEFM) – and, where possible, to assess CARE’s contribution to progress towards policy implementation goals. This executive summary provides a brief overview of the report.

The Gender Scorecard

Tool Overview. The Gender Scorecard for U.S. Administration Officials is a systematic approach to tracking the extent to which officials engage in actions aligned with CARE’s policy implementation goals on GBV and CEFM. This scorecard grew out of CARE and APEP’s earlier work co-developing the Champion Scorecard, a tool designed to assess whether Members of Congress who participated in CARE’s Learning Tours became stronger champions of legislation aligned with CARE’s foreign assistance policy goals. The Gender Scorecard adapts the Champion Scorecard so that, rather than focusing on the actions of those responsible for passing legislation, the tool captures information relevant to assessing the actions of those responsible for implementing policy – that is, administration officials.



WHAT: The scorecard consists of 29 actions that would indicate the extent to which an official is serving as a “champion” for GBV or CEFM issues within the administration and helping to move key pieces of policy implementation forward. Each action is assigned a numerical score and organized into three categories corresponding to different levels of “champion-ness” (demonstrates interest, promotes awareness and understanding, and promotes implementation and funding of policy).

WHO: We tracked 12 officials in the White House, State Department, and USAID.

HOW: We gathered data on officials’ actions from 2013-2016, using LexisNexis searches of public records, CARE’s internal database for tracking its interactions with officials, and searches of government websites and social media pages.

How the Scorecard Was Useful. CARE staff report that the scorecard helped CARE be systematic in tracking progress over time and revealed strategically useful insights for focusing CARE’s advocacy. For example, it gave CARE a much better understanding of the targeted officials’ roles and which spaces they were active in, helping CARE focus its advocacy efforts more effectively. In addition, the process of tool development was valuable for clarifying what policy implementation looks like and how to influence it.

Limitations of the Scorecard. A key challenge was that administration officials are often hard to track due to lack of information on their actions and unpredictable turnover. APEP and CARE found it especially difficult to track funding, which is a key indicator of policy implementation. In addition, the scorecard was primarily useful for monitoring the extent to which administration officials engaged in actions consistent with CARE’s policy goals; it was less well-suited for assessing CARE’s contribution to changes in champion scores.

The Quality of Discourse (QOD) Content Analysis

Tool Overview. The Quality of Discourse tool was designed as a complement to the Gender Scorecard and emerged in part to address the Scorecard’s limitations. While the scorecard tracks whether an official said something, it does not track the “quality” of what was said, that is, the extent to which the official’s discourse reflected CARE’s advocacy messages. CARE promotes a more nuanced understanding of the complex causes and consequences of CEFM, and potential strategies for addressing it. Using a content analysis approach, we developed a Quality of Discourse tool to assess the extent to which officials’ public statements reflected the nuanced understanding of CEFM that CARE advocates.

WHAT: We developed a content analysis codebook consisting of 33 codes, each corresponding to specific issues and solutions that CARE identified as central to its advocacy on CEFM.

WHO: We included largely the same officials we tracked in the scorecard.

HOW: We gathered all pieces of public discourse that we could find from 2013-2016 (e.g., speeches, testimony, op-eds, social media posts), using the same data collection processes that we used for the scorecard. We then analyzed each piece of discourse to determine the extent to which it reflected CARE’s key messages, as captured by the codes in the codebook.

How the QOD Was Useful. The QOD provided a systematic basis for understanding where discourse started and how it changed over time, offering useful insights to CARE’s advocates about how to adjust their messages. Because the codebook was designed to closely map onto messaging that was largely unique to CARE and a few coalition partners, the QOD findings also provided evidence of the contribution that CARE and its allies were making to changes in discourse. In addition, the process of developing the codebook and aligning it with CARE’s core advocacy messages was a useful exercise, providing an opportunity for CARE to specify the language its advocates were using and wanted to see echoed in officials’ discourse.

Limitations of the QOD. There were gaps in our ability to gather social media data, and we were able to gather only limited insights into the quality of social media discourse. This is partly a reflection of the limited resources we could devote to this component of the work, and our need for additional technical expertise in social media analysis in particular. Another limitation: QOD findings alone don’t allow us to say whether positive changes in public discourse translate into positive changes in programmatic or funding priorities. The scorecard, which is designed to track actions not words, offers potentially complementary insights here.

Lessons and Recommendations for Future Applications

Ensure appropriate resources and internal processes are in place. We discovered that the tools required considerable labor, as well as clear data collection processes. Make sure there is sufficient organizational and staff capacity to collect and analyze the data, and set clear expectations around “who, what, when and how” – who is responsible for collecting or reporting what, when it needs to happen, and how to do it.

Think ahead: how will you use the data? When piloting the tools, we made a fundamental mistake by not sufficiently considering how the data would be used. Our suggestion: think about which data will tell you what you truly need to know. This will simplify the process of determining what to collect, and help prevent a lot of data collection and analysis that is minimally useful.

Strategically time your reports. We learned that more frequent or strategically timed updates would have improved CARE’s capacity to fully utilize the findings. When planning the reporting schedule, think about when you expect to achieve different markers of progress. At what time points would it be most strategically useful for you to analyze the data and take stock of progress? Thinking strategically about the timing of reports can make the data more useful, enabling advocates to more clearly see the meaning of the findings and their connection to the strategy.

Carefully weigh whom to track. It can be difficult to determine which officials are most important to track. Some useful questions to help guide the selection process include: Who is in a position to make key decisions? Is that person someone whom you can influence? Does it make sense to track an individual or an entire office? Which levels in the hierarchy should you track? Can you collect adequate data on a targeted official?

Tailor the tools to help identify your contribution. Whether you are thinking about adapting the scorecard or trying out a QOD content analysis, we suggest tailoring the tools to match the kinds of things you can realistically influence. What changes in discourse would signal you have helped influence the narrative? What actions might an official take that would indicate your advocacy is having its intended impact on policy implementation? Calling on you for information? Releasing a strategy that contains specific sections you advocated? Funding a specific type of program?

Consider the context: Is QOD right for you? Our pilot of the QOD content analysis focused on a relatively narrow issue area (child marriage) and CARE was one of relatively few organizations pushing these specific messages and nuances. This allowed us to use the QOD analysis to learn whether CARE was making progress in its influencing work, and to help detect CARE’s contribution to any changes in the quality of discourse. The tool would have yielded fewer useful insights if CARE had been working in a large coalition of organizations all saying the same thing, or had focused on a very large and crowded issue area. Consider whether you are working on a broad or narrow issue, how crowded the advocacy field is, and what you want to learn. The QOD content analysis might be most helpful in cases where: the issue domain is relatively small; the field is sparse; and/or your message is unique or you are introducing a new idea/narrative.

Help crack the tough nuts: funding and social media. Tracking funding and social media are two areas where APEP and CARE fell short in piloting these tools. We encourage others to continue seeking better ways to capture useful data in these two areas – and sharing lessons learned. For example, we imagine there are many productive ways to leverage the information on social media – e.g., by tracking the quality of hyperlinked content, or examining how an advocate’s messages are spread through mentions by influentials. Our preliminary efforts to analyze grant opportunities posted on grants.gov¹ suggest that this may be a promising source of information on whether and how the government funds gender equality and female empowerment programs. Future efforts along these lines could offer insights into government funding priorities and the ways that targeted issues are incorporated into funding opportunities.

CARE and APEP developed these tools in the spirit of experimentation. We wanted to see what was possible and what we could learn. We hope the lessons we learned from piloting these tools are useful to others seeking to understand and assess progress toward their vision of a better world.

¹ Grants.gov is a central portal where organizations and individuals can electronically find and apply for grants throughout the federal government.

I. INTRODUCTION

From 2009 to 2016, CARE USA and the Aspen Planning and Evaluation Program (APEP) partnered to develop a set of tools for evaluating policy advocacy. Our aim was to track progress toward CARE’s policy advocacy goals, assess CARE’s influence, and provide insights to inform CARE’s strategy. Along the way, we learned helpful lessons about how to approach advocacy evaluation, and what kinds of opportunities – and challenges – evaluators and advocates face in collecting reliable, useful, and meaningful data.

In this report we share what CARE and APEP learned from our experience piloting two tools focused on CARE’s advocacy on gender issues in particular: the Gender Scorecard for U.S. administration Officials, and the Quality of Discourse content analysis. We developed and piloted these tools from 2013-2016.

CARE and APEP developed these tools in the spirit of experimentation. We wanted to see what was possible and what we could learn, with the expectation that we would sometimes get it wrong. Advocacy evaluation was – and still is – a relatively new field. Over the past decade, evaluators and advocates have developed new tools and approaches to try to address some of the unique challenges of measuring advocates’ contributions to complex policymaking processes. Further, assessing policy implementation offers specific challenges and has received less attention from evaluators. Experimenting with new approaches, and sharing what worked or did not work, can help the field continue to move forward. We hope that this report helps advocates, evaluators, and others in related fields build on the approaches we tried and the lessons we learned.

IN THIS REPORT

- The context for developing the tools
- An overview of the tools and their intended purpose
- Examples of the kinds of evaluative and strategic insights the tools can provide
- Lessons about the strengths and limitations of the tools
- Recommendations for future applications of the tools

II. THE CONTEXT

In 2013, CARE’s USG advocacy on gender equality issues was focused primarily on two significant policies that had recently been adopted: the **United States Strategy to Prevent and Respond to Gender-Based Violence Globally** (“US GBV Strategy”), an inter-agency strategy released by President Obama in 2012 that aims to integrate and elevate GBV prevention and response into all relevant areas of U.S. foreign policy; and the **Violence Against Women Reauthorization Act (VAWA) of 2013** adopted by Congress and signed into law in early 2013. While VAWA is focused on addressing GBV in the US domestic context, the reauthorization included a provision that “The Secretary of State shall establish and implement a multi-year, multi-sectoral strategy to prevent child marriage.”² Prior to the adoption of these policies, CARE had engaged in significant advocacy efforts for a number of years to raise attention to issues related to GBV and child marriage globally and to encourage increased U.S. Government action. These advocacy efforts were

² See <https://www.congress.gov/bill/113th-congress/senate-bill/47/text>.

conducted in partnership with coalitions such as Girls Not Brides (GNB) USA and multiple other coalitions and individual organizations.³ CARE and its partners formulated and implemented long-term advocacy strategies that involved actively working with Congress and the administration and engaging CARE’s grassroots supporters to increase support for action through coordinated advocacy and communications efforts.

With these two key policies in place, CARE’s advocacy to promote greater integration of girls’ and women’s empowerment and gender equality into the U.S. Government’s foreign assistance programs focused on two goals:

1. Implementation of the U.S. inter-agency GBV Strategy, including provision of adequate funding and transparent reporting on implementation;
2. Release of a comprehensive inter-agency strategy to prevent child, early, and forced marriage (CEFM), and implementation of this strategy, including provision of adequate funding and transparent reporting on implementation.

These goals signaled a shift for CARE’s advocacy on GBV and CEFM: from pushing for adoption of legislation to pushing for implementation of policies – and consequently, from focusing on Members of Congress to focusing on administration officials responsible for implementing policy. This shift prompted CARE to reflect on what it means to do effective advocacy around policy implementation, as opposed to passage of a bill:

- What happens after laws are passed – how are policies implemented once they are adopted?
- What kinds of measures and data collection are needed to track whether and how a policy is being implemented?
- What information would tell CARE that its advocacy efforts are making a difference in how policies are implemented?
- Who does CARE need to influence in policy implementation?

These questions guided CARE and APEP’s collaborative effort to develop and pilot two tools. First, we developed the Gender Scorecard for U.S. Administration Officials, a systematic approach to tracking the extent to which officials engage in actions aligned with CARE’s policy implementation goals on GBV and CEFM. This scorecard grew out of CARE and APEP’s earlier work co-developing the Champion Scorecard, a tool designed to assess whether Members of Congress who participated in CARE’s Learning Tours became stronger champions of legislation aligned with CARE’s foreign assistance policy goals.⁴ Using the Champion Scorecard as a starting point, CARE and APEP sought to adapt it so that, rather than focusing on the actions of those responsible for passing legislation, the tool would capture information relevant to assessing the actions of those responsible for implementing policy – that is, administration officials.

Over the course of piloting the Gender Scorecard from 2013-2016, we worked to refine the tool, making adjustments so that it better reflected champion actions that were both measurable and meaningful. We quickly discovered that the scorecard was not well-suited for evaluating a crucial goal of CARE’s advocacy: to influence how officials talk about CEFM. CARE promotes a more nuanced understanding of the complex causes and consequences of CEFM, and potential strategies for addressing it, drawing on insights from its child marriage programming globally.⁵ Although the scorecard tracked whether an official said something, it did not assess the “quality” of what was said, that is, the extent to which

the official’s discourse reflected CARE’s advocacy messages. To complement the insights we were able to glean from the scorecard, we developed a second tool to assess the quality of officials’ public discourse on CEFM using a content analysis approach.

We approached this process of developing the two tools with optimism tempered by modest expectations. We hoped to gather useful information on whether and how policies were being implemented, and perhaps even evidence of how CARE contributed to progress. But we recognized the difficulties inherent in attempting to monitor administration officials, as well as the challenges of identifying CARE’s contribution to a dynamic and complex policy implementation process. We also acknowledged our own resource limitations: we were working within an existing contract, and in the case of the quality of discourse work, we had limited resources to develop a tool that was not included in the original scope of work.



3 Key organizations partnering with CARE on CEFM and GBV advocacy include the International Center for Research on Women, International Women’s Health Coalition, Futures Without Violence, Vital Voices, Amnesty International and the American Jewish World Service.
4 For more information on the Champion Scorecard, see David Devlin-Foltz and Lisa Molinaro, “Champions and “Champion-ness”: Measuring Efforts to Create Champions for Policy Change” (Center for Evaluation Innovation: Washington, DC, 2010).
5 For more on CARE’s child marriage work, see www.caretippingpoint.org



III. THE GENDER SCORECARD FOR U.S. ADMINISTRATION OFFICIALS

What Did We Hope to Learn? - The Purpose of the Scorecard

As we set out to develop this tool, we had three expectations about the purpose it would serve.

- 1. Track policy implementation:** We expected that the scorecard would enable CARE to track the extent to which key administration officials took actions that promoted or signaled progress toward policy implementation (e.g., prioritization of GBV and CEFM, more funding, more political support). As part of this, CARE hoped the scorecard would help shed light on key questions around where funding goes, how it is being used and what is “new” money versus reallocated money.
- 2. Gather monitoring and evaluation (M&E) lessons through experimentation:** We wanted to gather insights that would strengthen our collective understanding of how to monitor policy implementation and to evaluate related advocacy. What data could we reliably gather? Would we be able to say anything useful with the data we collected? Would measurable things be meaningful, and would meaningful things be measurable? What kinds of resources and capacities are needed?
- 3. Help clarify CARE’s strategic thinking:** The third expectation was that the process of developing the scorecard would help CARE clarify what “policy implementation” looks like (i.e., what actions would indicate that implementation is taking place), whom CARE’s advocacy targets should be, and which advocacy activities to prioritize.

Tool Overview

Identifying Measurable and Meaningful Actions to Track. APEP worked closely with CARE staff to define the kinds of actions that would indicate an official was serving as a “champion” for GBV or CEFM issues within the administration and helping to move key pieces of policy implementation forward. Each action falls into one of three major categories:

SCORECARD CATEGORY	DESCRIPTION	EXAMPLES
Demonstrating interest	Includes relatively low-effort activities that reflect an official’s inclination to learn more about GBV or CEFM.	Requesting information from CARE and its allies, or visiting one of CARE’s or its allies’ development projects.
Promoting awareness and understanding	Encompasses activities that show increasing commitment to raising awareness of GBV and CEFM issues.	Giving speeches, writing op-eds, and co-hosting events.
Promoting implementation and funding of policy	Focuses on higher-level implementation and funding activities.	Organizing stakeholder consultations, releasing reports on strategy implementation, and releasing new funding or programming related to GBV or CEFM.

CARE staff and APEP then collaborated to rank the actions, assigning a numerical score (1 to 8). The highest scores were reserved for actions in the third category of promoting implementation and funding of policy. The lowest scores were assigned to lower-effort actions dealing primarily with expressing interest in the issues. (See Appendix A for the full list of scorecard actions and their scores.)

We collected retrospective data on officials’ actions from one year prior to the time that CARE engaged them with targeted advocacy efforts.⁶ This was intended to provide a baseline against which to evaluate changes in their behavior over time. We also assessed whether the officials increasingly engaged in higher-scoring actions, signaling a greater effort or commitment to implementing policies in line with CARE’s objectives.

The point values assigned to each action – and the resulting overall cumulative score – are in some sense relative, depending on the starting point for each official. For an official whose starting point is minimal awareness of GBV or even opposition to CARE’s objectives, for example, simply getting that person to meet with CARE is a big win. For an official who is a longtime champion of GBV prevention, a meeting with CARE may be a less momentous event. As one CARE staff member put it: the journey of each individual is more important than the score itself. We focused more on movement and the nature of actions taken rather than on absolute scores.

Selecting Officials to Track. The State Department, USAID, and the White House were the three key bodies responsible for implementing the US GBV Strategy and Violence Against Women Reauthorization Act. In collaboration with its coalition partners, CARE did a power mapping of the three institutions to identify who could make relevant decisions on GBV and CEFM, and whom CARE could plausibly influence.

⁶ This varied, depending on when the official assumed office.

APEP and CARE piloted the scorecard with twelve officials in the State Department, USAID, and the White House. We intended to include actors at diverse points in the policy implementation process. For example, we included decision makers directly focused on issues of gender equality, women, and girls, as well as those focused on other interrelated issues (e.g., health, migration, human rights). We also sought to incorporate senior officials at the USAID Mission and Embassy in selected countries to capture policy implementation at the country level. We also decided to include the Secretary of State and the USAID Administrator – not because we expected CARE to directly influence them, but rather because their actions and statements on GBV and CEFM serve as important signals of agency priorities and activity. Over time, we modified this list to reflect lessons we learned as we went along; by the pilot’s conclusion, we were tracking ten officials.

Data Collection and Analysis. CARE and APEP both contributed to data collection for the scorecard. APEP used LexisNexis to monitor the public actions and statements of officials. LexisNexis is a subscription service that facilitates searches across thousands of major and industry-specific media outlets in English, including television news transcripts, Congressional transcripts, press releases, newspaper articles, published speeches, web blogs (previously vetted by LexisNexis), and others. Through LexisNexis, APEP set up a system to automatically run keyword searches and send us daily updates on any items that referenced an official’s name (or title) and a set of relevant terms (e.g., child marriage, child bride, married adolescents, gender-based violence). For retrospective searches, APEP ran these LexisNexis searches manually.

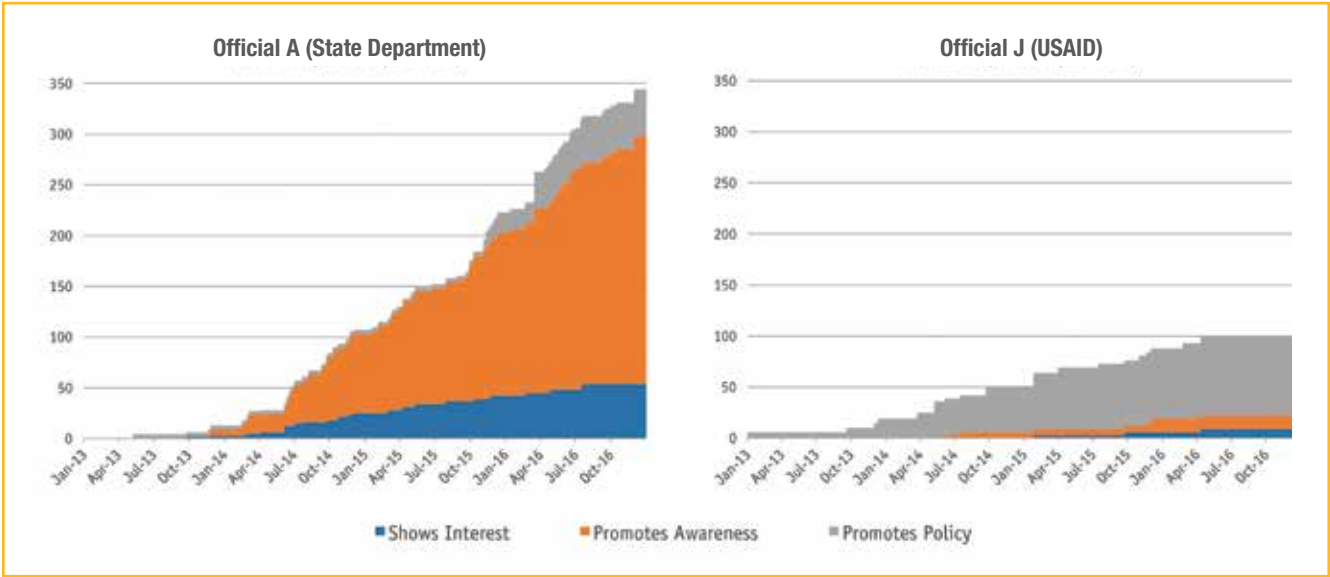
CARE used its internal database, CongressPlus, to track all interactions with officials, detailing whom CARE staff met with, the extent of the interaction, and additional notes on how the interaction went. CARE also monitored the social media accounts of the officials we were tracking, using a combination of Twitter’s advanced search function and manual searches of Twitter and Facebook pages. CARE sent all the information it collected to APEP on a monthly basis. In addition, CARE and APEP routinely monitored the websites of the State Department, USAID, and the White House. We searched for information on officials’ publications, press releases and announcements, and public remarks. This helped capture any items missed through LexisNexis.

APEP used Microsoft Excel to create a database of champion activity. For each action an official took, we recorded basic information (e.g., official name, date of action, keywords, action description) and coded the action according to the categories in the scorecard. To report on the results, we used formulas to generate a cumulative score for each official, which fed into a step chart showing how the cumulative score changed over time, broken down by scorecard category (see Figure 1). We also created a final scorecard table showing the number of times each action was taken and the cumulative score, by action, scorecard category, and overall (see Appendix A).

Examples of Scorecard Findings

Officials varied widely in the degree to which they engaged in champion activity. By systematically collecting information on officials’ actions over a four-year period, we aimed to evaluate the extent to which they exhibited champion behavior over time – as reflected by changes in their cumulative scores over time. A steady increase would suggest a constant rate of champion activity over time; a sharp change in the slope would suggest the official became a stronger champion; and a plateau would suggest lack of activity.

One of the ten officials we tracked exhibited a change in slope, suggesting that her activity intensified beginning in the fall of 2015. Four others showed fairly steady increases in their scores over time. For the remainder, we recorded only modest amounts of champion activity. Figure 1 shows sample area graphs.



Promoting awareness was the primary focus of most officials’ champion actions. Champion activity was particularly concentrated on “promoting awareness and understanding” (e.g., public speaking and writing), which is the middle category in the scorecard. In seven of the ten champion scorecards, actions that promote awareness made up the largest proportion of the total score. Official A in Figure 1 shows this quite clearly; in fact, 244 out of the 344 points that make up Official A’s total cumulative score are the result of actions promoting awareness. This finding partly reflects the specific responsibilities of different offices. For example, Official A in Figure 1 was a very public-facing official, whereas Official J was less public-facing but more involved in the details of policy development. Official J’s score, while lower than that of Official A, was almost completely comprised of actions in the highest scorecard category of policy implementation and funding – and in that way, these actions were qualitatively more significant.⁷

We observed evidence that some key steps of policy implementation were achieved. Despite the general preponderance of actions to promote awareness, by the end of 2016, we had recorded key markers of progress in the highest scorecard category: promoting implementation and funding of policies. These included actions that CARE and its allies were pushing for, such as: organizing consultations with NGOs and other stakeholders, releasing reports showing the implementation status of the US GBV strategy (an indicator of transparent reporting), and releasing the US Global Strategy to Empower Adolescent Girls.

⁷ The disproportionate number of actions to “promote awareness” observed in most champion scorecards may also be partly a methodological artifact: it is relatively easy to collect data on speeches, op-eds, and other actions in this middle scorecard category, especially compared to collecting data on actions in the highest scorecard category of promoting implementation and funding of policies.

How the Scorecard Was Useful

The scorecard helped us be systematic in tracking progress. The record of champion activity helped CARE take stock of progress made over time, providing a more systematic higher-level view than, for example, a narrative account of particular “wins” or milestones.

The scorecard helped reveal strategically useful insights for focusing CARE’s advocacy. The close monitoring of these targets gave CARE a much better understanding of the targeted officials’ roles, the inner workings of their offices, and which spaces they were active in. For example, the scorecard helped show which offices were or were not centrally involved in child marriage discussions, helping CARE focus its advocacy efforts more effectively. CARE also leveraged lessons we learned about the challenges of tracking funding (described below). Our experience attempting to collect funding data revealed that the government did not in this instance provide sufficiently detailed or accessible sources of information, prompting CARE to articulate a policy recommendation around clearer accounting of program funds and accountability for implementation.

The process of tool development was valuable for clarifying what policy implementation looks like and how to influence it. Developing the scorecard helped CARE unpack what it means to see policies implemented, particularly in the context of a complex bureaucracy like the State Department or USAID. The process of identifying actions to track in the scorecard helped CARE define specific markers that would signal progress toward policy implementation. It also pushed CARE to think critically about how best to focus its efforts and resources. For example, CARE defined specific aspects of work focused on child marriage, as opposed to GBV more broadly. Deciding what point value to assign each action in the scorecard also proved helpful in revealing the relative weight CARE placed on different aspects of cultivating champions and promoting policy implementation. For example, CARE realized that tracking funding – or trying to “follow the money” – was an important, if challenging, priority and would be valuable not just to CARE but to others in the advocacy and development community.

Limitations of the Scorecard

Administration officials are often hard to track due to lack of information on their actions and unpredictable turnover. We observed considerable variation across the officials we attempted to track: some had relatively high champion scores, while others were quite the opposite. While some of this variation likely reflected “real” differences in how much officials took champion actions, we suspect that it was also partly driven by our limited ability to collect data on the behavior of some of the officials. Indeed, a lack of data led us to abandon our initial efforts to track two officials at the Mission and Embassy level in specific countries; we simply could not collect sufficient information to say anything meaningful. Our LexisNexis searches, which had proven useful for tracking Members of Congress, were less effective for tracking administration officials whose actions were less likely to appear on the public record. Although CARE was able to provide useful behind-the-scenes insights into what was happening within the relevant government agencies, there were many discussions and meetings that non-government organizations like CARE were not part of. This made it hard to know how and why certain decisions were made. In short, the inner workings of government agencies are difficult to track.

An additional complication in data collection was turnover in key offices. For example, when a high-placed official we were tracking stepped down in 2015, there was a lag of several months before his successor was appointed, disrupting our ability to assess the extent to which the new individual in this office became a stronger champion over time.



Funding is an elusive target. Our efforts to collect information on funding reinforced how complex it is to answer two key questions among advocates: is this new money or is existing money just being moved around, and how is the money being used? Although the scorecard included actions specifically focused on funding, we were largely unable to provide satisfactory answers to these questions. The challenge was mostly a function of incomplete information. In order to glean insights into funding for GBV or child marriage programs, we relied primarily on public announcements of initiatives and on the government’s status reports on the implementation of the GBV Strategy. These were relatively imprecise and incomplete sources of information for determining how much funding was going to which specific programs. APEP’s interviews of well-placed Congressional staff revealed that they were also unable to get sufficient information about the sources and uses of funding to implement the GBV strategy.

The scorecard provided relatively few insights into CARE’s contribution. We expected it would be difficult to use the scorecard to identify CARE’s unique contribution, in part because CARE largely does advocacy in coalition with peer organizations. For most scorecard actions, it would be hard to draw a line directly from CARE’s advocacy activities to an official’s behavior. There are a few exceptions: for example, if we observed an increase in the frequency of an official’s interactions with CARE, or if an official tweeted about CARE’s gender and empowerment work, we might infer that CARE was at least informing that official’s thinking. But in general, we found that the scorecard was primarily useful for monitoring the extent to which administration officials engaged in actions consistent with CARE’s policy goals, rather than for assessing CARE’s contribution to changes in champion scores. This early finding in part led us to develop the quality of discourse tool to gather complementary insights into CARE’s potential influence.



IV. QUALITY OF DISCOURSE (QOD) CONTENT ANALYSIS

What Did We Hope to Learn? - The Purpose of the QOD Tool

CARE and APEP had three main expectations regarding what we would learn from the QOD content analysis:

- 1. Systematically understand what officials say and how that changes over time:** We expected that the QOD tool would enable CARE to track the extent to which the public statements of key administration officials conveyed a nuanced understanding of the complex causes, consequences, and solutions regarding CEFM – and whether the quality of these public statements improved over time. CARE drew a potential connection between officials’ public discourse and policy implementation. Because these officials were key actors in implementing U.S. policies on CEFM and GBV globally, their discourse could be one indicator of the government’s priorities. Their words could also help shape narratives around CEFM. For example, if decision makers in gender-focused offices began incorporating new nuances into their public statements about child marriage, this could serve as an early indicator of a shift in their thinking and priorities. Or if child marriage began to appear more frequently in very senior officials’ public discourse, this could serve as a broader signal of the agency’s priorities in policy implementation. In short, what made it into a public statement mattered.
- 2. Assess CARE’s influence:** A second expectation was that the QOD tool would help CARE track its influence on administration officials. By tailoring the tool to focus on the specific messages that CARE and its allies were advocating – including ones that CARE believed few or no other advocates were pushing – we aimed to help identify CARE’s contribution to any observed changes in discourse. Of note is that unlike the scorecard, the QOD tool was focused specifically on tracking discourse on CEFM, as opposed to statements related to GBV more broadly. The decision to focus on CEFM was because CARE’s messaging on this issue was more sharply drawn and distinct than its GBV messaging, which we believed would make it easier to assess if and whether CARE’s CEFM messages were having an influence in the discourse of advocacy targets.
- 3. Inform CARE’s strategy:** The third expectation was that the QOD findings would help inform CARE’s advocacy strategy by providing insights into which of CARE’s messages may have resonated more – or “gotten through” – and which may need tweaking or additional emphasis.

Tool Overview

Developing a Codebook to Define “Quality.” The centerpiece of our QOD content analysis was a codebook that defined the attributes of discourse we were tracking. The codebook reflected the core drivers of CEFM and recommended solutions identified through CARE’s program learning and evidence, including its 2016 research report, *The Cultural Context of Child Marriage in Nepal and Bangladesh: Findings from CARE’s Tipping Point Project Community Participatory Analysis* (the “Tipping Point report”).⁸ This report summarized some of the main messages that CARE’s advocates emphasized to administration officials and wanted to see echoed in officials’ discourse.

The codebook contains 33 codes, each corresponding to specific issues and solutions that CARE identified as central to its advocacy on CEFM. These codes fall into three main categories: drivers of CEFM; impacts and problems associated with CEFM; and action steps that policymakers, funders, nonprofits, and others should take to address CEFM. A single piece of discourse can contain multiple codes – e.g., if an official described multiple drivers of CEFM. In addition, we developed a coding rule to assess the degree to which a piece of discourse focused on CEFM. The purpose of this coding rule is to help distinguish between cases where CEFM was only mentioned in passing versus cases where champions focused more strongly on CEFM and connected it to the key issues and solutions for which CARE advocated.⁹ The codebook also includes two additional codes for identifying cases where officials mentioned the countries where CARE’s CEFM work was being connected to global advocacy most proactively: Bangladesh and Nepal. As a whole, the codebook is designed to allow us to examine the extent to which champions’ discourse on CEFM incorporated the points that CARE highlighted in its Tipping Point report and emphasized in its advocacy activities. (See Appendix B for a summary list of codes.)

Selecting Officials to Track. Because the QOD tool was developed to complement the scorecard, we tracked the same list of officials. The one exception was the then-First Lady of the United States, whom we added to the QOD analysis in 2016. Although the First Lady was not among the officials we were tracking in the scorecard, CARE had been engaging with her office for some time to increase her interest in taking on girls’ empowerment as part of her agenda. Because she is a highly influential voice, and had begun to mention CEFM more in her public discourse, CARE asked APEP to add her to the QOD analysis.

Data Collection and Analysis. The QOD content analysis included all public discourse on CEFM that we could find for each official. We separately collected and analyzed two forms of discourse: longer-form discourse (speeches, press releases, interviews, op-eds, blog posts, fact sheets, testimony) and social media discourse (tweets, Facebook posts). APEP and CARE gathered these pieces of discourse as part of the data collection processes described above for the scorecard – namely, through LexisNexis searches, monthly updates from CARE on officials’ activities, and searches of US government press office webpages, blogs, and social media pages. To be included in the content analysis, a piece of discourse had to mention child marriage, early marriage, forced marriage, or child brides at least in passing. As noted above, the discourse analysis was focused specifically on child marriage, not GBV.

To streamline qualitative data storage and analysis, APEP used a qualitative data analyses web application called Dedoose.¹⁰ APEP uploaded the full text of all discourse pieces to Dedoose, and then applied codes to relevant excerpts within each discourse piece. Once all discourse was coded, Dedoose allowed us to easily identify the frequency with which different codes were used, and showed precisely where and to what extent different officials mentioned specific points that CARE advocates.

⁸ See *The Cultural Context of Child Marriage in Nepal and Bangladesh: Findings from CARE’s Tipping Point Project Community Participatory Analysis* CARE 2016

⁹ The coding rule allowed us to code each piece of discourse as having a weak, moderate, or strong focus on CEFM, based on the extent to which the champion’s reference to CEFM incorporated “codable” discourse – i.e., discourse that mentioned an issue or solution contained in the codebook. The coding rule was defined as follows: Weak focus on CEFM – CEFM was mentioned briefly, and was the focus of one or no codable sentences; Moderate focus on CEFM – CEFM was mentioned briefly, and was the focus of two or more codable sentences; Strong focus on CEFM – CEFM was discussed as a primary focus of at least one paragraph. For example, if an official mentioned child marriage in passing while discussing women’s empowerment, without mentioning any of the key points that CARE advocated, this would be an example of a “weak” focus on CEFM.

¹⁰ Dedoose is a cloud-based software that allows the user to create individual codes corresponding to a codebook; upload data in different formats; code excerpts within a longer piece of text; apply descriptors to an entire piece of text; and sort, analyze, and export the data.

Examples of QOD Findings

LONGER-FORM DISCOURSE

The volume of longer-form discourse mentioning CEFM increased over time. As shown in Table 1, officials mentioned CEFM in their public discourse much more in 2016 compared to 2015, 2014, and especially 2013. Table 1 also illustrates that four of the eleven officials provided the bulk of the longer-form discourse on CEFM.¹¹

Table 1. Pieces of Discourse Mentioning CEFM, by Official and Year

CHAMPION	2013	2014	2015	2016	TOTAL
Official A (State)	1	9	14	13	37
Official B (White House)	0	4	5	13	22
Official C (State)	1	5	2	5	13
Official D (USAID)	0	1	4	5	10
Official E (State)	1	1	0	3	5
Official F (White House)	0	0	0	2	2
Official G (USAID)	0	0	0	1	1
Official H (State)	0	0	0	1	1
Official I (USAID)	0	0	1	0	1
Official J (USAID)	0	0	0	0	0
Official K (USAID)	0	0	0	0	0
TOTAL	3	20	26	43	92

Much of the longer-form discourse only weakly focused on CEFM. A majority of discourse items (65%) were coded as having a weak focus on CEFM. For example, one high-placed State Department official, who wrote an op-ed to mark International Women’s Day in 2015, briefly mentioned early and forced marriage, but did not refer to any of the

¹¹ The lack of longer-form discourse in 2013 likely partly reflects changes in government priorities and staffing; for example, key officials we tracked came into their positions in late 2013 or early 2014. It may also be partly a methodological artifact; we began tracking discourse on an ongoing basis in 2014, and ran retrospective searches for discourse in 2013. Our ability to retrospectively locate some discourse from 2013 may have been limited by factors like “link rot” (links pointing to pages that are no longer available) and removal of previously posted material.



issues or solutions included in the codebook. This suggests that in many cases, officials mentioned CEFM in passing, but failed to discuss it in a way that reflected the nuanced understanding CARE advocated. The remainder of discourse had either a moderate (22%) or a strong (13%) focus on CEFM. A “strong” focus means that CEFM was discussed as a primary focus of at least one paragraph (examples in Figure 2). Four champions had examples of discourse with a strong focus on CEFM.

There was wide variation in how often different codes appeared in longer-form discourse. Figure 3 shows examples of codes and the relative frequency with which each code appeared in officials’ discourse. The four codes at the top clearly stood out as the most frequently mentioned. At the other end of the spectrum are codes that were mentioned in few or no pieces of discourse. This helps reveal which nuances officials were – or were not – emphasizing in their public discourse on CEFM. For example, although champions frequently cited the challenges faced by married girls, they almost never mentioned support for programs and policies tailored to the needs of already married adolescents.

In addition, from CARE’s perspective, the significance of how often a code was mentioned depended on the specific code. For example, observing frequent references to girls education was not particularly ground-breaking, whereas even a single mention of access to sexual and reproductive health and rights services for girls, comprehensive sexuality education for girls, or the link between child marriage and controlling sexual purity was significant, given the political climate and constraints around what officials would or would not say publicly about these issues.

Figure 2. Examples of Discourse with a Strong Focus on CEFM

- USAID official’s testimony before the Senate Committee on International Operations and Organizations, Human Rights, and Democracy.
- State Department official’s keynote address at the Gender 360 Summit.
- State Department official’s remarks at the Launch of UNFPA’s State of the World Population Report.

Figure 3. Assessing Quality: Which Nuances Were – or Were Not – Mentioned in Longer-form Discourse

OFFICIALS MOST FREQUENTLY MENTIONED...	# DISCOURSE PIECES
Discrimination against girls and the low value of girls.	34
Complex barriers to education for girls.	31
Expanding and promoting access to formal and non-formal quality education for girls.	27
Challenges faced by married girls.	19
OFFICIALS SOMETIMES MENTIONED...	
Provide access to sexual and reproductive health and rights services for girls.	10
Causes of child marriage in situations of crisis and conflict.	10
Promote and fund multi-sectoral strategies that focus on CEFM, and CEFM programs that are integrated into other broader programs (e.g., education, health).	9
Strengthen implementation of minimum age of marriage laws and policies.	7
Provide comprehensive sexuality education for girls.	6
The link between family honor, sexual purity, and controlling sexuality.	5
Include men and boys in efforts to address CEFM.	5
OFFICIALS NEVER MENTIONED...	
Promote positive role models of gender equitable roles and relationships.	2
Fund and support programs and policies tailored to the needs of already married adolescents.	1
Geographical, seasonal, and environmental factors (including migration).	0
Impacts of child marriage on boys.	0
Stigma of divorce.	0
Invest in youth-led organizing efforts.	0

*Note: figure shows a selection of the 33 codes in the QOD codebook.

Officials’ longer-form discourse exhibited signs of positive change over time. There was evidence of a positive trend in the frequency with which officials mentioned some of the issues and solutions in the QOD codebook. Of particular significance, certain nuances around sexuality and dowry, which CARE and its allies were unique in emphasizing, began to appear more frequently in officials’ longer-form discourse. None of the officials mentioned the link between family honor, sexual purity, and controlling sexuality in 2013 and 2014; but in 2015, a key official mentioned this nuance twice, and in 2016, three officials did so. The frequency with which officials mentioned the need to provide access to sexual and reproductive health and rights services also showed a steady increase between 2013 and 2016.

In addition, two key officials publicly mentioned dowry for the first time in 2016. These changes provide some evidence of the contribution that advocacy efforts by CARE and its allies made in encouraging more nuanced public discourse on CEFM.

SOCIAL MEDIA DISCOURSE

Patterns in social media discourse diverged from longer-form discourse. Between 2013 and 2016, we recorded a total of 137 social media items mentioning CEFM. One official stood out for a particularly high frequency of mentioning CEFM on social media, relative to the others we tracked; her posts made up 50% of all social media discourse we collected, even though she contributed only about 11% of the longer-form discourse on CEFM. By contrast, two other officials who mentioned CEFM relatively frequently in their longer-form discourse only mentioned this issue once each on social media in the four-year period, despite being active social media users. This underscores differences across officials in terms of where – and how – they raise the issue of CEFM.

We also observed that social media discourse rarely mentioned CEFM as a form of violence or in connection with Let Girls Learn, a major US government initiative on girls’ education. Officials only occasionally mentioned CEFM in the same post as Let Girls Learn or GBV. This differs from longer-form discourse on CEFM, which frequently referred to GBV and girls’ education.

Most social media posts contained insufficient information to code for quality, though some linked to more substantive content. Most posts and tweets conveyed too little information to be coded using the codebook. They made simple statements or highlighted events, publications, and other activities. Some posts did link to more substantive information intended to inform or raise awareness. A notable example of this is one official’s tweet linking to CARE’s 2016 Tipping Point report. Figure 4 provides examples of social media posts that contained sufficient information to code for quality.

FIGURE 4. EXAMPLES OF SOCIAL MEDIA POSTS CONSISTENT WITH QOD CODES



State Department Official: The Syrian refugee crisis has increased #childmarriage due to displacement & poverty #GBVchat #MENA #CEFM



State Department Official: What are the key drivers of #childmarriage in Nepal & Bangladesh? Find out in CARE's new Tipping Point report here: <http://bit.ly/1XKCDqF>



State Department Official: "Concept linking #sexuality. #child marriage, #honor is about perceived protection rather than actual protection." @CARE @PlanGlobal #CSW60



State Department Official: Earlier today in Copenhagen I interviewed Tinbit Daniel, a young woman who uses media to show girls different role models. I asked Tinbit about her vision for 2030. She said she hopes no girls are married at age 13 and instead, that girls are visible, valuable, and powerful. #WD2016

How the QOD Content Analysis Was Useful

The process of developing the codebook clarified the meaning of "quality" discourse. Developing the codebook helped CARE clarify for itself what constitutes discourse that reflects CARE's understanding of the nuances of CEFM and the strategies needed to address it. The process of aligning the QOD codes to CARE's core advocacy messages provided a useful opportunity for CARE to specify the language its advocates were using and wanted to see echoed in officials' discourse.

The QOD provided a systematic basis for understanding where discourse started and how it changed over time. By systematically tracking and coding four years' worth of discourse, we were able to develop a clearer picture of whether and how officials were talking about CEFM, and how their discourse changed over time. The QOD provided CARE with insights into which nuances officials seemed to be picking up, which nuances remained absent, and whether there was any evidence of improvements in their public discourse.

Tracking discourse offered (at least indirect) insights into the government's priorities. For advocates, knowing that an official began mentioning CEFM or particular nuances more frequently in public discourse served as one signal regarding the priorities of that office.

The QOD findings helped CARE identify its influence. CARE found the QOD findings useful for distinguishing its influence on changes in discourse over time. This was partly because the codebook was designed to closely map onto CARE's specific messaging on CEFM, and partly because there were relatively few other actors (to CARE's knowledge) emphasizing some of these messages. So when an official began incorporating into public remarks one of the unique messages CARE was strongly pushing, we could conclude with some confidence that CARE had likely contributed to that shift in discourse. Even in cases where CARE was part of a coalition of organizations emphasizing certain messages, having a systematic record of discourse – showing that these messages begin to appear – helped provide evidence of CARE's and its peers' contribution.

The QOD findings helped advocates adjust their messaging. CARE's advocates found the QOD tool intuitive and appealing, as it reflected the nuanced messaging they used to try to deepen officials' understanding of very complex issues. Advocates felt they could interpret the QOD findings about which messages were or were not appearing in officials' discourse and draw conclusions about the results of their efforts, and about which messages were resonating. For example, if officials began talking more about dowry, an advocate might pick up on this more in discussions with them. Or if officials were not talking about the role of engaging men and boys, an advocate might think about whether she was providing the information or explanation that officials needed.

Limitations of the QOD Content Analysis

Public discourse does not necessarily signal action. In CARE's view, it matters what officials say publicly about CEFM. The individuals we tracked were all key spokespeople for the U.S. government, and influential voices at the national and international level. A shift in their discourse toward greater nuance can help set the tone and direction of policy conversations, or suggest a shift in their programmatic or funding priorities. However, just because an official begins mentioning CEFM more – or becomes more nuanced – we cannot infer that the official will take action aligned with CARE's policy objectives. Here the scorecard, which is designed to track such actions, offers potentially complementary insights.

There were gaps in our ability to gather social media data. We ran into problems of data availability when collecting data on Twitter and Facebook activity, particularly retrospectively. For example, Twitter's advanced search function does not cover all tweets; full access to historical data requires a considerable fee and technical expertise. Facebook search options were limited, making data collection time-consuming.

We gathered only limited insights into the quality of social media discourse. Our efforts to track the quantity and quality of social media discourse on CEFM did not yield very useful findings. This is perhaps partly a reflection of the limited resources we could devote to this component of the work, and our need for additional technical expertise in social media analysis in particular. We found that most tweets and Facebook posts do not contain sufficient information to code for quality. This is a data point in and of itself, albeit a basic one: most social media discourse does not meet CARE's standards of "quality" as captured by the codebook. But we suspect that more useful insights about social media discourse could have been derived if we had approached the analysis differently. For example, our analysis did not extend to hyperlinked content. Resources to engage technical expertise, and careful forethought regarding how to make sense of social media data, could yield more useful insights into the quality of social media discourse.



V. LESSONS AND RECOMMENDATIONS FOR FUTURE APPLICATIONS

Our experience piloting the two tools yielded many lessons that CARE and APEP will apply to future evaluation and advocacy work. Below we offer some key takeaways and recommendations for how others might consider using or adapting the tools in the future.

Ensure Appropriate Resources and Processes Are in Place

Lesson learned: The two tools required considerable labor and clear data collection processes. Because this was an experiment in understanding how to track policy implementation within the Administration, we were “learning while doing.” We learned that data collection for the scorecard and QOD content analysis required a considerable time investment. For example, CARE had to track every meeting its staff attended and record what was said, adding additional work for staff juggling many demands on their time. There was a learning curve around developing internal processes to ensure that CARE staff reported when they had meetings and the content of those meetings. Gathering the social media data was also very time-intensive. CARE designated a Fellow whose primary responsibility was helping to collect data for the scorecard and QOD. CARE also developed documentation for how to collect and record the data so that incoming staff could learn the processes. On APEP’s side, the process of reviewing the data CARE collected, cross-checking it against LexisNexis hits, and coding it for the scorecard and QOD content analysis was quite labor intensive.

Recommendation: Make sure there is sufficient organizational and staff capacity to collect and analyze the data. Set and reinforce clear expectations around “who, what, when and how” – who is responsible for collecting or reporting what, when it needs to happen, and how to do it.

Think Ahead: How Will You Use the Data?

Lesson Learned: We put too much emphasis on producing data, rather than consuming data. This is a common problem: focusing on collecting as much data as possible, and then struggling to analyze it all in a timely way. We might have simplified our task of figuring out what to track, and how, if we had more strictly focused on the question of “what is the most important information to have,” either for informing CARE’s strategic planning or assessing its contribution. In short, APEP and CARE did not focus enough on the end user, or consumer, of the data we gathered.

Recommendation: Think about how you will use the data – as much as how you will collect it. Which data will tell you what you truly need to know? This will simplify the process of determining what to collect, and help prevent a lot of data collection and analysis that is minimally useful.

Strategically Time Your Reports

Lesson learned: More frequent or strategically timed updates would have improved CARE’s capacity to fully utilize the findings. Because this pilot was folded into an existing contract, we reported our findings from the scorecard and QOD pilots as part of larger annual reports encompassing other related advocacy evaluation activities. Although CARE did use these annual reports – both for reporting to funders and to inform its own internal strategic discussions – it provided only a single reflection point for the entire year. For CARE’s staff, reading this kind of lengthy report at the end of the year was daunting. It was difficult to find time to dig into the findings and fully utilize valuable insights. Given additional resources, we expect it would have been useful to provide more strategically timed reports to advocates on the Gender Scorecard and QOD findings.

Recommendation: When planning the reporting schedule, think about when you expect to achieve different markers of progress. For example, consider whether the actions you are tracking in the scorecard include those you might expect an official to do more frequently or readily, as well as those that may take longer to achieve or happen infrequently. At what time points would it be most strategically useful for you to analyze the data and take stock of progress? Thinking strategically about the timing of reports can make the data more useful, enabling advocates to more clearly see the meaning of the findings and their connection to the strategy. (Of course, sometimes the reporting schedule is determined by funders or other considerations as well.)

Carefully Weigh Whom to Track

Lesson learned: There are many factors to consider in deciding whom to track. Through our experience piloting the two tools, we discovered that selecting individuals or entities to track involves a certain calculus. We learned that the particular decisions we made early on around whom to track had implications – both positive and negative – for the information we were able to collect and the conclusions we could draw. Although there is not a single “right” answer to the question of whom to track, we offer some recommendations below about considerations that may help guide the selection process.

Recommendations:

- **Who is in a position to make key decisions, and whom can you influence?** The list of officials we tracked included those with whom CARE’s advocates directly interacted, as well as a few very high-level individuals whom CARE was unlikely to directly influence. We included the latter to gain a clearer understanding of what these key actors were saying and doing, but were more cautious in drawing any connection between CARE’s advocacy and these officials’ actions and discourse.

- **Does it make sense to track an individual or an office?** An advocate typically exerts influence through her relationship with the individual official. But successfully changing how that individual thinks or acts is useless if he or she leaves the critical office in which decisions are made. On the other hand, tracking an office (regardless of who holds the position) may make it harder to identify an advocate’s influence. When a new person arrives, the advocate may be starting a new relationship from scratch. When deciding whether to monitor an individual or his/her office, it may be helpful to consider how long the official is likely to hold that position.
- **Which levels in the hierarchy should you track?** For example, it might be helpful to include representatives of different levels, and with varying responsibilities and issue expertise, to get a more complete picture of government priorities. If you are working on a smaller issue, such as child marriage, it can be important to know when high-level officials comment on it. You might want to also track a key mid-level official with more issue-specific expertise to gather information on how nuanced the agency’s understanding of the issue is.
- **Can you collect adequate data on a targeted official?** The scorecard and QOD content analysis are most useful for more public-facing officials, and perhaps for officials about whom you can gather useful behind-the-scenes information.
- **Should an official’s subordinates “count”?** For example, if an advocate meets with an official’s deputy, that might be counted as a meeting with the official because we can reliably say that the deputy influences the official’s thinking. This works best for small, relatively self-contained offices; for a sprawling agency like USAID, it would be difficult to say who is communicating with or influencing whom.
- **How might you build flexibility into the selection process?** In large measure, we tracked the same individuals across the entire piloting period. But it may be helpful to allow for greater flexibility in responding to advocates’ engagement of new potential champions when they emerge on the scene. For example, you might pick a few “evergreen” champions, whom you know are in positions of influence and will likely speak on your target issues. And then leave a few open spots for more “opportunistic” targets that emerge and can be incorporated into the tracking process retroactively.

Tailor the Tools to Help Identify Your Contribution

Lesson learned: As noted above, the scorecard – at least in the form we piloted – was not particularly well-suited for identifying CARE’s contribution to progress. We included various actions that, while certainly an important part of developing policy champions and implementing policies, could not easily be connected back to CARE. If the scorecard had included a more “niche” set of actions that CARE was advocating, this may have yielded more insights into CARE’s influence. For the QOD content analysis, we focused more narrowly on the specific (and where possible, unique) messages CARE was advocating, which made the tool more useful for evaluating CARE’s influence.

Recommendation: Whether you are thinking about adapting the scorecard or trying out a QOD content analysis, we suggest tailoring the tools to match the kinds of things you can realistically influence. What changes in discourse would signal you have helped influence the narrative? What actions might an official take that would indicate your advocacy is having its intended impact on policy implementation? Calling on you for information? Releasing a strategy that contains specific sections you advocated? Funding a specific type of program?



Consider the Context: Is QOD Right for You?

Lesson learned: In our pilot of the QOD content analysis, we were focused on a relatively narrow issue area – child marriage – and CARE was one of relatively few organizations pushing specific messages and nuances around this issue. This allowed us to use the QOD analysis to help us learn whether CARE was making progress in its influencing work, and to help detect CARE’s contribution to any changes in the quality of discourse. But the QOD analysis would have yielded less useful insights if CARE had been working in a large coalition of organizations all saying the same thing, or had focused on a very large and crowded issue area.

Recommendation: Consider whether you are working on a broad or narrow issue, how crowded the advocacy field is, and what you want to learn. If you are interested in gathering insights into your potential influence on discourse, the QOD content analysis might be most helpful in cases where:

- The issue domain is relatively small;
- The field is sparse; and/or
- Your message is unique or you are introducing a new idea/narrative.

A QOD content analysis could be used in other ways. For example, you might wish to gather systematic descriptive data on the quality of discourse in a given field or among a group of actors as part of formative research to inform a message strategy. In that case, considerations like issue size or a crowded field may matter less.

Help Crack the Tough Nuts: Social Media and Funding

Lesson learned: Funding and social media are two areas where we fell short in piloting these tools. Our initial attempt to track funding, relying on announcements of funding and CARE’s behind-the-scenes insights, was largely unsuccessful. And our foray into tracking social media was relatively limited, constrained in part by resources and time. We encourage others to continue seeking better ways to capture useful data on funding and social media – and sharing back lessons learned. This would be a valuable contribution to the fields of advocacy, development, and evaluation.

Recommendation for Leveraging Social Media: We imagine there are many productive ways to more fully leverage information on social media. For example, we did not track the quality of hyperlinked content, which may provide insights into the kinds of organizations, ideas, initiatives, and messages that officials are promoting to their vast social media networks. Another potential opportunity would be to use data on CARE’s own social media pages to assess how its messages are spread through mentions by influentials.

Recommendation for Tracking Funding: We believe there is value in trying to track funding, given its centrality to the policy implementation process. Our efforts to track funding for GBV and CEFM programming are situated within the broader field of gender responsive budgeting, which focuses on ensuring public budgets reflect a commitment to gender equality.¹²

One additional potential source of information on funding is grants.gov, which lists grant opportunities for USAID and the Department of State, among other federal agencies. In 2016, APEP piloted an approach to tracking Requests for Applications (RFAs), using a search of grants.gov to identify all funding announcements or requests for applications (RFAs) since 2013 that incorporated CEFM or GBV. The pilot revealed some evidence of funding opportunities that we had not captured through existing data collection efforts, and provided some preliminary qualitative insights into how RFAs incorporated CEFM (e.g., as part of results indicators vs. a gender analysis). A limitation of this approach is that the data speak to funding opportunities not the ultimate amount of funding granted. It also requires time to review individual RFAs, and a clear plan for how to track and make sense of the data.

From a strategic learning standpoint, we expect that this approach would be most useful if data were collected longitudinally so that changes in the quantity and quality of funding opportunities could be assessed. It might be helpful to further hone the focus to specific countries the U.S. Government has made a priority for a given issue, and track RFAs for those countries in particular. Future efforts along these lines could offer insights into government funding priorities and the ways in which targeted issues are incorporated into funding opportunities.

VII. CONCLUSION

We set out to experiment with new ways to measure progress toward policy implementation and CARE’s contribution to that process. Some of our efforts were more successful than others, but all yielded useful learning that will inform our future advocacy and evaluation work. We hope the lessons we learned from piloting these tools are useful to others seeking to understand and assess progress toward their vision of a better world.

¹² See, for example: <http://www.unwomen.org/en/digital-library/publications/2010/1/introduction-to-gender-responsive-budgeting-and-aid-effectiveness> and <https://www.usaid.gov/kosovo/news-information/press-releases/launching-gender-responsive-budgeting-kosovo>.



APPENDIX A: GENDER SCORECARD FOR ADMINISTRATION OFFICIALS

CATEGORY	ACTION	BASE SCORE	FREQUENCY	OVERALL SCORE
1. Demonstrates Interest in GBV (inclusive of Child Marriage)				
Events	Attended an event related to GBV or child marriage.	1	0	0
	Attended a CARE event.	2	0	0
Travel	Visited development projects of CARE and/or its allied organizations that address GBV or child marriage.	3	0	0
Information Seeking	Requested information (via phone, email, etc.) from CARE and/or its allied organizations related to GBV or child marriage.	3	0	0
	Held in-person meeting with CARE and/or allied organizations, including key coalitions such as Girls Not Brides US and the GBV coalition.	3	0	0
SUBTOTAL				0

Gender Scorecard for US Administration Officials (continued)

CATEGORY	ACTION	BASE SCORE	FREQUENCY	OVERALL SCORE
2. Promotes Awareness and Understanding of GBV (inclusive of Child Marriage)				
Information Sharing	Invited peers, friends, or colleagues to learn about the objectives of CARE and/or allied organizations regarding GBV or child marriage.	2	0	0
Media	Wrote a blog post about GBV or child marriage consistent with the objectives of CARE and/or its allied organizations to internal website (i.e. own organisation's blog).	2	0	0
	Mentions CARE and/or related hashtags (e.g., #CARE) and either GBV or child marriage on Twitter or Facebook.	3	0	0
	Was interviewed in the media and made comments about GBV or child marriage consistent with the objectives of CARE and/or allied organizations.	3	0	0
	Wrote and placed op-ed about GBV or child marriage in an external publication (i.e. not own organisation's blog) consistent with the objectives of CARE and/or allied organizations.	3	0	0
	Issued press release about GBV or child marriage consistent with the objectives of CARE and/or its allied organizations.	4	0	0
Research	Contributed to evidence base about GBV or child marriage by conducting research or developing other credible evidence to support the policy objectives of CARE and/or allied organizations.	3	0	0
Public speaking	Spoke favorably about issue related to GBV or child marriage.	3	0	0
	Spoke at a CARE event related to GBV or child marriage.	4	0	0
Travel	Recruited colleagues to visit development projects related to GBV or child marriage consistent with the objectives of CARE and/or allied organizations.	3	0	0
	(Co)organized a trip to visit development projects related to GBV or child marriage.	4	0	0
Events	(Co)hosted events related to GBV or child marriage.	4	0	0
Partnerships	Formed new partnerships or alliances that support work addressing GBV or child marriage.	5	0	0
SUBTOTAL			0	

Gender Scorecard for US Administration Officials (continued)

CATEGORY	ACTION	BASE SCORE	FREQUENCY	OVERALL SCORE
3. Promotes Implementation and Funding of Policies				
a) US Strategy to Prevent and Respond to GBV Globally				
Coordination	Participated in an inter-agency forum meeting or mechanism aimed at coordinating USG agencies' efforts around GBV (i.e. implementing new strategy).	4	0	0
	Organized or participated in consultations with more than one organization (representing civil society, other governments, donors or other stakeholders) consistent with the objectives of the GBV strategy.	4	0	0
Monitoring, Evaluation, Learning & Reporting	Shared new data, learning, or research on GBV within or outside the agency.	4	0	0
	Released report showing status of implementation of strategy.	5	0	0
Integration	Engaged in or supported staff training or capacity building around GBV.	5	0	0
	Included GBV as a focus of key documents including operational plans, guidance documents and calls for proposals/funding.	5	0	0
Increased Funding or Scale-Up of Programming	Released new funding or programming related to GBV.	6	0	0
b) Special Focus on Child Marriage				
Child Marriage Strategy	Developed goals, objectives or indicators specific to child marriage in GBV strategy and/or adolescent girl strategy.	6	0	0
	Released comprehensive inter-agency child marriage strategy, as mandated by VAWA, Public Law 113-4 (Adolescent Girl Strategy).	8	0	0
Monitoring, Evaluation, Learning & Reporting	Released report showing status of implementation of strategy.	5	0	0
Funding and Scale-Up	Released new stand-alone child marriage projects and/or integration of child marriage objectives or indicators into other types of projects.	7	0	0
SUBTOTAL			0	
SCORECARD TOTAL			0	



APPENDIX B: SUMMARY OF CONTENT ANALYSIS CODES

Below we provide a summary list of codes used in the quality of discourse content analysis. The full codebook contains more detailed information and examples as guidance for coders.

Drivers of CEFM:

1. The link between family honor, sexual purity, and controlling sexuality
2. Dowry
3. Causes of child marriage in situations of crisis and conflict
4. Geographical, seasonal, and environmental factors
5. Lack of enforcement of child marriage laws
6. The inter-relatedness of multiple drivers of CEFM
7. Complex barriers to education for girls
8. Discrimination against girls and low value of girls
9. Exclusion of girls' voices in decision making, including in marriage

Impacts and Problems Associated with CEFM:

1. Impacts of child marriage on boys
2. Lack of access to SRH info
3. Stigma of divorce
4. Challenges faced by married girls

Actions Steps:

Access to services and opportunities

1. Provide access to sexual and reproductive health and rights services for girls.
2. Provide comprehensive sexuality education for girls.
3. Fund and support programs and policies tailored to the needs of already married adolescents.
4. Invest in infrastructure that improves mobility and access for adolescents, especially girls, in isolated communities.
5. Expand the choices available to young adults, especially girls, after secondary school by creating diverse ways to connect young people from remote and marginalized communities to vocational, livelihoods, and employment opportunities and markets.
6. Expand and promote access to formal and non-formal quality education for girls.
7. Support development of media strategies and technologies that can reach isolated communities and marginalized individuals within them (as a key way to portray alternative options for girls and foster debate around social norms and gender inequality).

LAWS AND POLICIES

1. Eliminate laws that permit marriages of children with parental or judicial consent or make no reference to the right to choice in marriage.
2. Strengthen implementation of minimum age of marriage laws and policies.
3. Go beyond legal age of marriage: strengthen policy interventions aimed at shifting gender norms and empowering adolescents to access social, economic and educational services and opportunities.
4. Strengthen citizenship, birth, and marriage registration systems, with a special focus on equitable access for girls.
5. Incorporate child marriage as part of planning/response from the onset of crises.

BIG PICTURE STRATEGIES

- 1. Give strong support and funding to local civil society and NGOs to take on and scale up community-based campaigns on CEFM.
- 2. Promote and fund multi-sectoral strategies that focus on CEFM, and CEFM programs that are integrated into other broader programs (e.g., education, health).
- 3. Champion girls’ rights and a focus on root causes of child marriage in multilateral and bilateral dialogues, particularly around M&E indicators.

INVOLVEMENT OF COMMUNITY MEMBERS

- 1. Include men and boys in efforts to address CEFM.
- 2. Create opportunities for community dialogue and awareness raising on gender equity and rights.
- 3. Invest in youth-led organizing efforts.
- 4. Promote positive role models of gender equitable roles and relationships.
- 5. Fund and support the training of community-based government workers in specific marginalized communities.

DEGREE OF FOCUS ON CEFM:

Each piece of discourse will be coded as having a weak, moderate, or strong focus on CEFM.

- 1. Weak: CEFM is mentioned briefly, and is the focus of one or no codeable sentences
- 2. Moderate: CEFM is mentioned briefly, and is the focus of two or more codeable sentences
- 3. Strong: CEFM is discussed as a primary focus of at least one paragraph

CARE’S FOCAL COUNTRIES:

This code captures whether a piece of discourse mentions one of CARE’s focal countries: Bangladesh, Nepal.



www.care.org

CARE USA

151 Ellis Street NE
Atlanta, GA 30303
USA
www.care.org

Founded in 1945 with the creation of the CARE Package, CARE is a leading humanitarian organization fighting global poverty. CARE places special focus on working alongside poor girls and women because, equipped with the proper resources, they have the power to lift whole families and entire communities out of poverty. Last year CARE worked in 87 countries and reached 82 million people around the world. To learn more, visit www.care.org.

The Aspen Institute is an educational and policy studies organization based in Washington, DC. The Aspen Planning and Evaluation Program, a consulting practice within the Institute, partners with foundations, nonprofit organizations, and individual funders to help them plan, evaluate, and strengthen their efforts to bring about positive changes in society. Learn more at aspeninstitute.org/apec.

PHOTO CREDITS: Cover: © 2015 Carey Wagner/CARE; Page 1: © 2017 Josh Estey/CARE; Page 2: © 2016 Carey Wagner/CARE; Page 7: © 2017 Elias Williams/IWHC; Page 8: © 2017 Elias Williams/IWHC; Page 13: © 2017 Elias Williams/IWHC; Page 14: © 2016 Carey Wagner/CARE; Page 17: © 2016 Carey Wagner/CARE; Page 22: © 2017 Allison O'Brien/CARE; Page 25: © 2017 Josh Estey/CARE; Page 27: © 2016 Carey Wagner/CARE; Page 30: © 2017 Elias Williams/IWHC