

Annex 4: On Nigeria Evaluation and Learning Framework (Design and Methods)

Introduction and Overview

Design: The evidence presented in this report fall under the On Nigeria evaluation and learning framework. This framework uses a mixed-methods, sequential design to measure progress toward outcomes and impacts at regular time intervals, and a combination of exploratory and descriptive designs to answer landscape and feedback evaluation questions. The evaluation employs complexity-aware approaches, which are appropriate given the theory of change whose causal relationships are dynamic, non-linear, and not always known at the outset. The framework includes 110 unique measures to track interim and long-term outcomes and impacts, as well as 64 questions pertaining to On Nigeria’s landscape (context, windows of opportunity, and theory of change assumptions) and feedback on implementation.

Data Sources: In total, the framework uses seven data sources. Primary source data include a national telephone survey, qualitative interviews and focus groups, media monitoring (including three data subtypes focused on conventional media, investigative reporting, and social media), and feedback workshops with grantees. Secondary source data include document review, grantee data, and global corruption indices. **Exhibit A4-1** presents the sampling technique, unit(s), and size for each data source.

Exhibit A4-1: Sample design and target sample size for primary source data and document review

Method	Sampling Technique	Unit(s)	2016	2017	2018	2019
 National Telephone Survey	Stratified random	Citizens	Not conducted	Not conducted	8,043	10,901
 Qualitative Interviews and Focus Groups	Purposive	Interview and focus group events	Not conducted	Not conducted	169	193
 Media Monitoring						
<i>Conventional</i>	Purposive	Media outlets	24	26	26	26
<i>Quality of Investigative Reporting</i>	Census	Articles/stories	1,266	Not conducted	1,461	Not conducted
<i>Social Media</i>	Purposive	Twitter hashtags, handles, and keywords	Not conducted	Not conducted	204	204
 Document Review	Purposive	Documents		450 (January 2015–August 2018)		390 (September 2018–November 2019)
 Grantee Data	Census	Grantee annual reports		28 (January 2016–August 2018)		52 (September 2018–November 2019)
	Varied	Other grantee reports and documents		Included in document review sample above		135 grantee-authored documents
	Varied	Grantee monitoring data		Not available		83 indicators from 39 grantees (2018–2019)
 Corruption Indices	Varied	Three global indicators: 1) World Bank Control of Corruption Indicator (index) 2) Gallup (poll) 3) Afrobarometer (poll)		Covering 2015–2017		Not available Not available

Method	Sampling Technique	Unit(s)	2016	2017	2018	2019
 Feedback Workshops	Census	Grantees	Not applicable ¹	Not applicable ²	55	72

Data Collection Timing: The national telephone survey and qualitative interviews and focus groups both consist of two cross-sections, having been collected once in 2018 and once in 2019; thus, they represent two “snapshots” in time. The 2018 telephone survey and qualitative questionnaires both asked respondents to make comparisons to prior years to obtain additional baseline insights. Media monitoring data are available back to 2016 through retrospective data collection and currently cover 3.5 years. Secondary source data cover documents authored from 2016 on, as well as select documents and indicators published in 2015 where useful to informing On Nigeria’s baseline.

Analysis and Synthesis: A strength of the mixed-methods approach is the ability to triangulate findings across a rich variety of data sources. This improves validity by ensuring findings are grounded in multiple perspectives, and providing opportunities to explore the complexity and nuance of findings. To bring together the diverse data, in both 2018 and 2019 the evaluation and learning partner first analyzed each data source separately. In each year, the team then synthesized findings over 2-day data analysis, triangulation, interpretation, and synthesis sessions for each module and the overall strategy to tell the story of On Nigeria to date. The evaluation and learning partner facilitated two rounds of multi-day participatory findings workshops with the On Nigeria Program Team and grantees to validate findings through feedback and input of additional data, as well as generate evidence-based suggestions for conclusions and considerations, which the evaluation and learning partner used as inputs when developing the final conclusions and considerations contained in this report.

Limitations: Data collection, sampling, and analysis have been designed to maximize evaluation rigor within the time and resources allocated. However, On Nigeria’s Evaluation and Learning Framework operates within four main design challenges:

- **Measuring corruption concretely and objectively:** Due to its nature as an illegal and generally hidden activity, and because corruption is a collective term covering a range of actions, measuring corruption directly is highly challenging and experts continue to debate the most accurate techniques. On Nigeria’s evaluation framework uses perceptions of corruption, direct experiences with bribery, citizens’ views of social norms related to corruption, and aggregate indices; use of transparency and accountability tools serve as partial proxy measures. However, each of these methods has its own limitations: corruption perceptions are susceptible to rapid shifts based on current events and extent of reporting itself may shift as a society confronts corruption; experiences with bribery capture just one of many types of

¹ The 2016 grantee workshop focused on generating early inputs to the theory of change.

² The 2017 grantee workshop focused on validating the theory of change and evaluation and learning framework.

corruption On Nigeria seeks to address; social norms are endogenously related to corruption (both driving it and resulting from it), but distinct from the acts themselves; aggregate indices are long-term trailing indicators that are insensitive to incremental changes; and use of transparency and accountability tools is a proxy that measures anticorruption actions rather than corruption itself and does not always have a strong empirical link to reducing corruption. To mitigate this challenge, the evaluation and learning framework uses multiple measures and data sources triangulated and analyzed within the broader political and economic context.

- **Analyzing On Nigeria’s contribution within a complex system and across various geographies in Nigeria:** The On Nigeria strategy involves numerous grantees seeking to influence various core components of the corruption system. Activities within each module attempt to change practices for a wide variety of actors: citizens, civil society organizations (CSOs), government actors, and the private sector; these activities and anticipated changes occur on independent timelines. Other actors’ anticorruption activities could also influence outcomes. Actual causes of observed outcomes and impacts cannot be determined with certainty, but landscape data capturing broader trends in the On Nigeria context help build a more robust understanding of the MacArthur Foundation’s specific role in promoting change through contribution analysis.

On Nigeria’s sectoral modules work in specific communities in specific states, while criminal justice, media and journalism, and cross-cutting grantees work more broadly. However, grantees working within a module are not necessarily working in the same community or state. The geographic diversity of On Nigeria’s programming complicates before/after analysis by increasing chances of differential effects across regions. Sampling designs for media monitoring and the national telephone survey help mitigate this limitation by ensuring broad representation of On Nigeria’s target regions.

- **Limited availability of public anticorruption monitoring data and response bias:** There are very limited secondary data available on most outcomes and impacts in the On Nigeria theory of change; the Nigerian state anticorruption agencies publish some data, but methods and publication timing are often not consistent. As a result, the evaluation and learning partner team has in many cases reconstructed public data based on press releases, grantee sources, and direct contact with governmental bodies.

All data collection with human subject respondents has potential biases arising from the cultural, socioeconomic, educational, ethnic, gender, and political backgrounds of data collectors and respondents. Careful training of data collectors, vetting and pilot-testing data collection tools, and effective probing help mitigate bias and response error among respondents. The national telephone survey, the most cost-effective method, samples among those with phones and necessarily underrepresents households without phones; weighting and estimation techniques help mitigate this bias.

- **Lack of baseline data for some measures:** The prior limitation regarding unavailability of historical secondary source data, combined with the fact that the first rounds of primary data collection occurred only in 2018 (almost 2 years after the end of the overall strategy

baseline), means that concrete baseline data are not available for most measures. Qualitative interviews and the national telephone survey included some questions to gather *reconstructed* baseline data for a limited number of outcomes, but these data provide only a perception of momentum, not a definitive status prior to On Nigeria’s launch. Only media monitoring data, which sample 2016, and some secondary sources (including corruption indices) provide quantitative measures that can be considered baseline. As a result of this data availability limitation, extent of progress may be underestimated if it began to occur prior to 2018 data collection.

Triangulation of data across various sources is intended to mitigate the limitations listed above, but it cannot remove the limitations.

The sub-sections below describe the design and sample, data collection method and tool(s), and limitations for each data source in detail. For additional detail on 2018 methods, see *Big Bet On Nigeria: 2018 Synthesis Report*.



National Telephone Survey 2018 and 2019

Design and Sample: The national telephone survey used a questionnaire administered via phone to a representative sample of Nigerian citizens to measure population-level changes in citizens’: (1) actions—seeking redress, demanding accountability, and engagement with social media; (2) perceptions of the extent of corruption; (3) attitudes and social norms surrounding corruption; (4) levels of trust in the government to tackle corruption; and (5) experiences with service delivery and media consumption in On Nigeria target sectors. In 2018, the survey was conducted nationally using a 35-item questionnaire. In 2019, the survey used a 28-item questionnaire that focused on the electricity sector, media and journalism, and general corruption among respondents nationally, as well as an education-specific 19-item questionnaire among parents of children in government primary or junior secondary schools both nationally and in On Nigeria target states (Kaduna, Lagos, and Ogun). Question wording, administration methods, and the sampling design were consistent across 2018 and 2019 to allow comparison, though some questions were added in 2019, which were not included in 2018. The sampling design for the education-specific poll in 2019 differs, however small sample sizes for these questions in 2018 precluded comparison of these data across the two years independent of the changes to the sampling design for this component.

EnCompass subcontracted NOIPolls (NOI) to administer the survey, with NOI’s database of 70 million phone-owning Nigerians serving as the sampling frame. NOI-assigned geographic quotas ensured that each state and senatorial district were proportionately represented in the sample. From this frame, data collectors completed 8,043 telephone interviews nationally in 2018. In 2019, data collectors completed 5,067 interviews for the national non-education survey; 3,327 interviews for the national education survey; and 3,045 interviews for the subnational education survey, which included 1,023 respondents from Kaduna, 1,013 respondents from Lagos, and 1,009 respondents

from Ogun. There are 538 overlapping respondents between the national and subnational education surveys in the three target states.

NOI and EnCompass constructed sampling (design) weights and post-stratification weights for all 2018 data and 2019 non-education data presented in this report, with post-stratification weights based on the 2006 Nigerian census. The sampling weight calculation accounts for the probability of selecting a respondent in each senatorial district, population coverage corrections, and non-response corrections, with the final weight for each respondent calculated as the product of the sampling weight and the post-stratification weight:

$$W(\text{respondent}) = \left(\frac{N_{1i}}{EA_i} \cdot \frac{N_i}{N_{1i}} \cdot \frac{EA_i}{AA_i} \right) \times PSW_j$$

Where:

N_i = Nigerian population age 18 and older (projected from 2006 census to 2017)

N_{1i} = NOIPolls number database age 18 and older in 2017 (sampling frame)

EA_i = expected sample in each senatorial district

AA_i = achieved sample in each senatorial district

PSW_j = post-stratification weight that adjusts for state and gender representation in final sample

For the 2019 education survey, EnCompass and NOI were not able to construct sampling (design) weights or post-stratification weights because there was no available population data on respondents' eligibility criteria (e.g., adults over 18 who have at least one child in government primary or junior secondary schools). NOI and EnCompass were only able to weight data based on non-response and over-response. Consequently, data for the 2019 education survey do not include confidence intervals and do not reflect population estimates; analysis primarily used descriptive statistics (as compared to inferential statistics for other sections of the survey).

In the final 2018 sample of 8,043 individuals, 42 percent were female, 70 percent were between 26 and 45 years of age, 53 percent had a post-secondary education, 69 percent lived in urban areas, and 50 percent had a monthly income of 60,000 Naira or less. Forty-eight (48) percent of interviews were conducted in English. Of these 8,043 respondents, 93 percent received their electricity from a DISCO, 30 percent had at least one child in a government primary or junior secondary school (thus, in a school eligible for UBEC funds), and 22 percent had at least one child in government primary or junior secondary schools and lived in a state with an operative HGSF program.

For the 2019 national survey related to electricity, media and journalism, and corruption in general, 48 percent were female, 55 percent were between 26 and 45 years of age, 53 percent had a post-secondary education, 71 percent lived in urban area, and 50 percent had a monthly income of 60,000 Naira or less. Forty-three (43) percent of interviews were conducted in English, and 93 percent of respondents received their electricity from a DISCO.

In the 2019 national sample for the education survey, 42 percent of respondents were female, 72 percent were between 26 and 45 years of age, 45 percent had a post-secondary education, 67 percent lived in urban areas, and 55 percent had a monthly income of 60,000 Naira or less. Thirty-five (35) percent of interviews were conducted in English.

For the 2019 education survey sample in the target states of Kaduna, Lagos, and Ogun, 44 percent of respondents were female, 67 percent were between 26 and 45 years of age, 39 percent had a post-secondary education, 70 percent lived in urban areas, and 49 percent had a monthly income of 60,000 Naira or less. Thirty (30) percent of interviews were conducted in English.

Data Collection Method and Tool: EnCompass developed the survey tool with input from a team of experienced Nigerian evaluators, corruption measurement experts, and NOI. Where appropriate, the team used validated survey instruments from Afrobarometer, the Global Corruption Barometer, and Corruption Victimization Surveys to inform question construction. In lieu of a baseline, one question asked respondents to compare the current prevalence of corruption (at both the national and state levels) to the prevalence 12 months prior, reflecting the direction of current trends. EnCompass updated the questionnaire in 2019 to capture data for new measures, remove baseline questions that had been answered in 2018, and further nuance existing questions.

NOI translated the finalized tools from English into four other languages (Hausa, Igbo, Pidgin English, and Yoruba). EnCompass consultants who are fluent in the respective languages and with experience collecting data on corruption issues, but had not previously been exposed to the English version of the questionnaire, back translated the language versions into English. EnCompass and this team then worked with NOI to reconcile differences and make final edits to translated versions. In 2018, two members of the EnCompass team attended enumerator training where further, slight modifications were made to the questionnaire for clarity based on feedback from enumerators—each fluent in the language they would administer the tool in. In July 2018, NOI piloted the original tool in all five languages with 543 respondents and analyzed the data with EnCompass; pilot data analysis looked for outliers in response by language to identify any potential questions where translations might be inaccurate, and concluded there were no outliers that had not shown similar cross-language variance in other surveys, such as Afrobarometer. EnCompass and NOI used qualitative feedback from enumerators, collected through daily pilot debriefs, to make final adjustments to phrasing. Pilot data were not included in the data presented in this report. In 2019, NOI pre-tested the questionnaires internally and conducted an abbreviated pilot of 100 respondents drawn nationally and across all five languages.

In 2018, NOI conducted full data collection from August 27 to September 28, with an average length of survey administration of 15.65 minutes. EnCompass conducted weekly data checks to ensure quality. Similarly, in 2019, NOI conducted data collection for the national non-education questionnaire from July 29 to August 16; for the national education questionnaire between August 19 and August 29; and for the subnational education questionnaire between August 19 and September 19.

Analysis: Data analysis used descriptive and inferential statistics to produce population-level estimates, expressed as a point estimate within a range reflecting the 95-percent confidence interval. Most data were disaggregated by On Nigeria target state and DISCO catchment area, as appropriate and feasible given sample sizes. A few questions—primarily those related to media and journalism and corruption more broadly—were further disaggregated by age, education, income, and geopolitical zone. Where estimates between 2018 and 2019 appeared to differ substantially but confidence intervals overlapped, statistical testing was conducted to determine whether there were statistically significant differences between the two years. All disaggregations presented in this report are statistically significant unless otherwise noted. Analysis was conducted in the Stata 14 software.

Limitations: The national telephone survey has the following limitations:

- Telephone administration excludes individuals without phones. If the demographics or responses of phone-owning Nigerians differ from those without phones, data would be biased toward the responses of those with phones. As described in the next point, comparison with household surveys administered face-to-face showed similar results for common items, but this cannot be inferred to mean there would be no differences on other items between Nigerians with and without phones.
- Key strategy-level “anchor” questions that align to Afrobarometer produced responses within each respective tool’s margin of error, indicating good cross-instrument reliability of these questions. However, this reliability cannot be assumed for other items, and it is not known whether phone or face-to-face survey administration would produce more accurate results.
- The 2018 sample size of 8,043 respondents was not sufficient to allow for statistically significant subnational disaggregation of questions related to the education programs (UBEC Intervention Fund and HGSF), which have restriction criteria that limit the number of eligible respondents.
- While the 2019 parent survey in the target states of Kaduna, Lagos, and Ogun was large enough to produce state-level estimates, the estimates could be slightly biased from true population estimates because the parent respondent pool might not be representative of the overall Nigerian population and EnCompass/NOI were unable to correct these estimates with survey weights because there is no population data available for these selection criteria in Nigeria.
- The sampling frame for the national education survey differed between 2018 and 2019, with the 2018 survey administered to a representative sample of adult Nigerians and the 2019 survey restricted to parents of students in government primary or junior secondary school. As such, national education data cannot be directly compared between 2018 and 2019.



Qualitative Interviews 2019

Design and Sample: The evaluation and learning partner conducted 193 qualitative interviews and focus groups in 2019³ to assess changes in respondents' awareness, experiences, and perspectives on anticorruption, transparency, and accountability issues since On Nigeria's inception. In collaboration with the MacArthur Foundation, the team designed a purposive sample that spanned seven main respondent groups to ensure perspectives were garnered across all modules and targeted states (Exhibits 1 and 2). The evaluation and learning partner increased penetration in targeted states for service delivery modules (HGSF, UBEC, and ELEC) in 2019 to enhance understanding of On Nigeria's effects and contribution to these modules in targeted states. This included increasing the number of LGAs and schools represented in the sample, and expanding respondent groups to incorporate perspectives from other supply chain and decision-making groups, such as vendors, distributors, cooks, and community leaders. Qualitative data collection in 2019 increased focus on Abuja FCT, and states adopting the ACJA to better achieve the sample for hard-to-reach respondent groups, including government watchdog organizations, and federal and state prosecutors and judges. The expanded sample breadth also included additional respondent groups in 2019 to address new module foci (i.e., party leaders, elected officials, social media experts, and behavior and norms change experts).

Data Collection Method, Tools, and Content: Qualitative interviews and focus groups captured (1) strength and momentum related to collaboration and anticorruption actions; (2) behavior change of key actors; and (3) perceptions of the most significant changes in systems and structures designed to reduce corruption. All qualitative interview and focus group discussions were conducted using MacArthur Foundation-approved semi-structured interview guides, and approved written or verbal informed consent from all respondents. Interviews employed Appreciative Inquiry, an asset-based approach that surfaces strengths and the most significant changes respondents identify, as well as a vision for the future and steps needed to make it a reality.

Data collection teams conducted two rounds of data collection. First, the evaluation and learning partner conducted eight grantee cohort focus groups directly with 51 grantees in Abuja in June 2019. The evaluation and learning partner then trained 12 data collectors with three teams of four conducting 185 interview events across nine states and Abuja/FCT (**Exhibit A4-2**) with non-grantees between July and August 2019. In nearly all cases, and unless respondents refused permission, data collectors recorded interviews and transcribed verbatim. In other cases, a trained transcriber took notes for the interviewer.

³ This represents a 14-percent increase over 2018, when EnCompass conducted 169 interviews and focus groups.

Exhibit A4-2: 2019 Qualitative sample: 193 interviews and focus groups (676 respondents)

Grantees / Sub-Grantees Cohorts	Non-Grantee CSOs and Parent Groups and Youth/ Women Influencers	Government Officials and Policymakers/ Legislators	Other Donors and Experts	Private Businesses (including vendors, aggregators, transmission companies and DISCOs)	School, LGEA, and SBMC Officials	Media Organizations and Journalists	Total
8 (70)	57 (64)	39 (42)	10 (11)	10 (11)	64 (459)	5 (19)	193

Analysis: A trained analysis team cleaned transcripts, developed a codebook, and coded and analyzed all transcripts using Dedoose and Excel software, applying coding techniques to allow for deductive and emergent thematic analysis. All transcripts were deductively coded to the On Nigeria measures, context, and assumption questions, while parallel inductive coding ensured the most common themes articulated by respondents surfaced.

Limitations: The 2019 sample oversamples On Nigeria’s target states, and reflects the 2019 goal of increasing education module penetration rates for qualitative data collection. Data analysts incorporated this understanding into analysis and interpretation processes.

Response bias could have been introduced from grantee focus group discussions. While including On Nigeria stakeholders was critical to capture perspectives on outcomes progress, and understand context and underlying assumptions, grantee respondents may have had a vested interest in focusing on successes. Data collectors employed probing techniques during interviews and corroborated data between respondent groups and other data sources to mitigate potential bias. Respondent group breadth came at the expense of depth for some modules (e.g., election, criminal justice, and behavior change modules), which limited the ability of the data to speak across levels of government, states, and stakeholder groups working within a given sector. To mitigate, data with insufficient corroboration from other sources were not included in findings. One of the objectives of data collection was to gain an understanding of awareness levels regarding mechanisms instituted to improve transparency and accountability from a range of respondent groups. Certain respondent groups lacked awareness, providing valuable insights. This lack of awareness could also affect responses to associated questions. Analysts mitigated the risk of incorporating false statements by including sufficient information during coding to assess the likely accuracy of one statement based on levels of awareness identified in related questions. The team also triangulated all qualitative interview data with other data sources, particularly document review, quantitative survey, and media monitoring.



Media Monitoring 2016—July 2019

Media monitoring involves three separate sub-streams of data collection and analysis: (1) *conventional* media monitoring, (2) quality of *investigative* journalism, and (3) *social* media monitoring.

Conventional Media Monitoring

Design and Sample: Media monitoring tracks reporting by conventional print, radio, television, and online media outlets over time, using a set of pre-identified keywords that correspond to On

Nigeria’s work in the modules and at the strategy level. This report presents data gathered retrospectively based on digital media archives for 2016, 2017, and 2018; ongoing media monitoring occurred in 2019 and will continue throughout On Nigeria’s period of activity. To ensure comparability of data over time, the methods align throughout all four years.

EnCompass subcontracted Playspread LLC, based in Lagos, to conduct the media monitoring. The media monitoring sample targets 24 media sources in 2016 and 26 sources from 2017 on, which EnCompass selected purposively in conjunction with Nigerian media experts at Playspread and the On Nigeria team (see **Exhibit A4-3**). The sample covers the most widely read and broadcast media sources in Nigeria, and media outlets that are On Nigeria grantees, and ensures regional representation of Nigeria’s main media markets. Although purposively drawn, the selection of media sources is designed to capture a broad swath of Nigerian journalism and be of sufficient size to reflect national trends in corruption reporting.

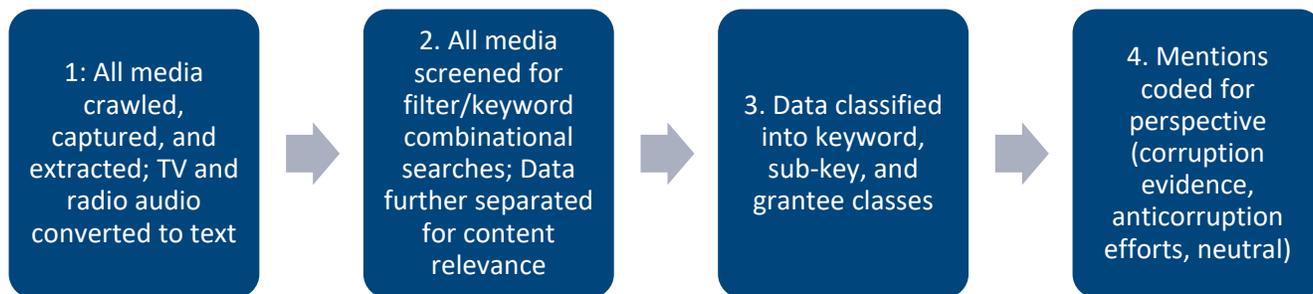
Exhibit A4-3: News sources sampled for media monitoring

Television	Radio	Print	Online
AIT Network	Raypower Network	This day	Daily Post
Galaxy Network (Lagos and Ibadan)	Brila FM (Abuja, Kaduna, Lagos, and Onitsha) (sports and news)	The Sun (sensational)	*The Cable
Channels Network (Abuja, Edo, Kano, Lagos)	Rhythm FM (Edo State, FRCN)	Vanguard (Edo State)	*International Centre for Investigative Reporting (ICIR)
NTA Network	KSMC Kaduna (FRCN)	Daily Times	*Premium Times
STV Network	Wazobia FM Lagos (in Pidgin English)	*Daily Trust (Abuja/FCT state/Kaduna state)	*Sahara Reporters
	Rock City FM (FRCN, Ogun State)	Leadership (Abuja/FCT state/Kaduna state)	
	Cool FM Abuja	Punch (Lagos State)	
		Guardian (Lagos State)	
		Nation (Ogun State)	

Note: * denotes a grantee news source; “Media name” denotes that this source was added after the 2016 baseline.

Data Extraction Method and Content: Media monitoring measures the level of coverage related to corruption issues and anticorruption actions, including the degree to which civil society’s and citizens’ anticorruption work is amplified through media coverage. Data consist of (1) numerical frequencies of keyword mentions, which capture the quantity of corruption-related reporting on different topics; (2) the overall tone (perspective) of articles—whether articles are focused on instances of corruption or anticorruption wins; and (3) an assessment of the quality of investigative journalism (see below).

Exhibit A4-4: Media monitoring process



Keywords are specific to each module and cover grantees, “voice” and “teeth” activities, and goods and service delivery. **Exhibit A4-4** shows the steps Playspread uses to collect, screen, and analyze the data.

EnCompass, with input from Playspread and the On Nigeria team, selected corruption- and anticorruption-related keywords for each module (148 in 2016, 207 in 2017, 207 in 2018, and 214 in 2019); media monitoring screened for these keywords alongside 35 corruption-related filters. Additional granting necessitated the addition of cross-cutting keywords. The complete list of keywords is provided in Annex 3.

Relevant keywords, also referred to as “mentions,” are reported and analyzed on a quarterly basis. Media monitoring data are analyzed in Excel, using descriptive statistics—primarily counts (frequency distributions and cross-tabulations)—to capture trends over time.

Limitations: Some grantee media sources are not available for retroactive monitoring through digital archives. However, the retroactive media monitoring sample is still large enough to ensure strong coverage of the Nigerian media landscape.

Quality of Investigative Journalism

Design and Sample: The 2016 analysis of investigative reporting quality drew from the 1,266 print and online articles in the 2016 conventional media monitoring dataset. The 2018 analysis drew from all 1,281 stories in the 2018 conventional media monitoring dataset, plus 180 additional articles submitted by grantees as investigative reports; in total, this yielded 1,461 news stories included in the quality of investigative journalism analysis.

A professor in the School of Communications at Lagos State University led this analysis with assistance from graduate student researchers who served as coders, under the Playspread’s supervision and EnCompass’ overall direction. Before measuring report quality, the lead researcher determined whether or not each article could be classified as “investigative.” Articles had to meet two standards—in-depth and proactive, as defined below. If a report met both, it qualified as an investigative report.

- **In-depth:** Reporting to investigate a single topic in-depth (i.e., examination attempts to be systematic, thorough, or present more than one perspective), typically to “uncover corruption, review government policies or corporate houses, or draw attention to social, economic, political, or cultural trends.”
- **Proactive:** Proactive reporting gathers information that was not previously public; it is **not** simply a passive reaction/report on press releases, government announcements, or related content.

2016 Method and Tool: Originally, only if an investigative article assessed first passed the in-depth and proactive test, it would then be measured for five standards of investigative quality: (1) public interest, (2) report originality, (3) neutrality of investigation, (4) research quality, and (5) source variety. To develop standards for quality, EnCompass conducted a literature review of reports and guides detailing investigative reporting. Sources from this review included the United Nations Office on Drugs and Crimes (UNODC), the United Nations Educational, Scientific and Cultural Organization (UNESCO), the Center for International Media Assistance (CIMA), and International Centre for Investigative Reporting (ICIR), among others. The five domains for quality used in this analysis are based on common themes that emerged from the literature review (See Annex 3 for the full tool).

To ensure inter-coder reliability, the coders received training and took part in inter-coder reliability tests conducted on a sample of 15 percent of articles. Based on coders’ Likert scale scores, the researcher and coders held discussions and determined a final set of scores for each article. EnCompass also audited and independently coded a random sample of 5 percent of the articles to ensure coder agreement.

Because few reports qualified as “investigative” according to the “in-depth” and “proactive” criteria, researchers applied the five standards to all 1,266 reports. The lowest score possible to assign was 5. Of the 1,266 articles, 1,242 received a score of 5. Overall, only six met investigative journalism quality standards, receiving a score of 17 or higher. Similar to traditional media monitoring, the assessment of the quality of investigative journalism also relied on frequencies; however, as the number of articles grows, it is expected that median will be used to measure central tendency.

2018 Method and Tool: In the course of the 2018 quality of investigative reporting process, data quality checks revealed a difference in the number of articles qualifying as investigative reports based on the coding conducted by the Lagos School of Communications team and the coding conducted by EnCompass. The Lagos School of Communications team identified 25 investigative articles in 2018, out of 1,461 total articles, while the EnCompass team’s coding indicated there could be 60 to 80 investigative reports among the entire article set.

To reconcile these differences, in the 2018 process, EnCompass tested the feasibility of a hybrid approach to the process. In the first step, EnCompass team members conducted an initial classification of all articles into six defined categories, including “investigative” and “potentially investigative.” An iterative process of reviewing and comparing results then helped further refine the methodology to clarify the definitions of investigative journalism and improve consistency across the EnCompass and Lagos Schools of Communications coding teams. As a next step to this process, the full set of articles classified as investigative and potentially investigative will be analyzed and scored on the five standards of investigative quality.

Limitations: Assessing the quality of an article requires coders to make a subjective judgment, which creates possibility that some degree of non-agreement can occur. However, an EnCompass audit found a 4.8-percent non-agreement rate, which is small enough to prevent a bias of overall results.

In designing this process, there exists little to no precedence for conducting an assessment of investigative reporting quality in the Nigerian media environment. This 2018 process, as well as subsequent data collection activities in future years, will serve as an important testing ground for developing a methodology that includes a thoughtful and concise categorization of Nigerian media reports, as well as a distinct set of investigative journalism quality standards.

Social Media Monitoring

Design and Sample: Social media monitoring tracks (1) the Nigerian social media landscape related to the On Nigeria strategy and modules, and (2) the social media activities of On Nigeria grantees to measure the level of citizens’ and civil society engagement in grantee-related and anticorruption conversations on Twitter.

This report presents data gathered retrospectively via the social listening platform, Pulsar, based on archived 2018 Twitter content. EnCompass chose to track data via Twitter rather than other social media platforms due to the ease of capturing broader quantitative conversation trends, as well as qualitative contextual information.

Data Extraction Method and Content: Using Pulsar, EnCompass overlaid filters and keywords (adapted from those used in conventional media monitoring) to identify the proportion of 2018 Twitter content originating from Nigeria that was corruption- and anticorruption-related.

On Nigeria strategy- and module-level monitoring captured quarterly totals of module-level tweets, total module engagement (numbers of “likes” and “re-tweets”), and assessments of qualitative content related to spikes in posts and conversations.

Grantee-level monitoring was able to identify total Twitter output and subsequent engagement per grantee over time, as well as the proportion of content produced per grantee directly related to corruption and anticorruption.

Limitations: Because this is a new process, EnCompass will continue to refine the 2018 social media monitoring methodology before moving to collect any future or retrospective data. Due to the various uses and possible contexts of many of the selected keywords and filters, EnCompass found during data quality assessments, that the qualitative content of tweets was not always relevant to On Nigeria strategy, modules, and corruption or anticorruption content. The main adjustments needed involve including more specificity in keyword sets, and piloting different combinations of keywords and filters to discover which yield the most relevant results.



Document Review 2016–2019

Design and Sample: Document review provides data for a variety of measures, including court records and legal cases; allocations of service-directed monies and expenditures; supply chain data; meter installment geographical distributions; and studies and reports from civil society, grantee annual reports, government, and the international community.

Document review covers relevant documents published from January 1, 2016 to November 2019. In limited instances, documents published in 2015 that were highly relevant to evaluation questions were also included.

Document review used a screening process to identify relevant documents during each round of data collection and analysis. Document collection compiled news items and reports the MacArthur Foundation and grantees forwarded with documents the evaluation and learning partner independently identified through searches. During the first round of document review (January 2015–August 2018), this first step yielded more than 600 documents. During the second round of document review (September 2018–November 2019), there were over 700 documents included in the first step. At the second stage, documents were screened for relevance; all those deemed relevant to On Nigeria measures or context and assumption questions were included in the review. Documents were drawn from a variety of sources, including the media, international and Nigerian nongovernmental organizations’ reports, donor reports, academia and think tank publications, grantee reports, workshop notes, presentations, and other relevant documents.

Data Extraction Method and Content: Documents varied widely in content. Consequently, the data resulting from this exercise include but are not limited to the following: (1) strength and momentum related to collaboration, capacity, and anticorruption actions; (2) behavior change of key actors,

including government, donors, grantees, and citizens; (3) evolution of corruption- and anticorruption-related laws and policies; (4) content containing key contextual information relevant to a particular module or the On Nigeria strategy; and (5) anticorruption-related activities and/or content grantees generated.

Analysis of January 2015–August 2018 Documents: After the initial scan, evaluation team members identified 478 relevant documents (including grantee reports), and assigned excerpts from these documents to one or more of the 102 codes related to On Nigeria theory of change measures, context questions, and assumptions across all modules. The team then conducted an analysis of the 1,367 coded excerpts, and produced a summary document related to each module.

Analysis of September 2018–November 2019 Documents: In the second round of analysis, the team identified 577 relevant documents (including grantee reports), and assigned relevant excerpts to one or more of 156 codes related to module measures, context questions, and assumptions. The module-level summary documents produced included an analysis of 2,859 coded excerpts.

Limitations: Throughout initial document collection, some documents were unavailable, or contained limited or incomplete information. Due to the volume of documents, limitations within this activity varied based on particular documents. Documents came from a variety of sources, including media, donors, and grantees; therefore, different documents could have particular biases, based on the authorizing source. Similarly, the volume of reports did not necessarily allow for a quality check on all data published in them, such as government or DISCO reports.



Grantee Data 2018 and 2019

Grantee data include available grantee annual reports, other reports and publications produced by grantees (e.g., reports, presentations, meeting minutes, guides, public statements, evaluations, assessments), data grantees collect as part of implementation activities (e.g., beneficiaries reached, training reports and statistics, public content produced), and data from grantees' own monitoring and evaluation activities (e.g., tracking, surveys, assessments). The grantee data validation exercise, initiated in summer 2018 with the technical assistance partner, verified what data grantees were actually collecting (whether for monitoring or as part of grant activities) and when, and helped identify 12 evaluation and learning measures for which grantee data were determined to be a priority source for the 2019 data collection. A small number of additional grantee data resources were identified at the November 2019 grantee convening.

Because grantee data are a secondary source, the evaluation and learning partner does not control the timeline of their availability. Consequently, this synthesis report only includes data that (1) grantees analyzed into a summary format (e.g., no raw data); (2) mentors and grantees shared with the evaluation and learning partner as of November 2019; and (3) the evaluation and learning partner determined to be of sufficient quality in method and collection to cite (e.g., reasonable sample sizes for the types of inferences made and valid tools).

This report cites six main pieces from the subset of grantee data that met these criteria:

1. *Grantee Annual Reports 2016, 2017, 2018, and 2019* (reviewed during document review process)
 - A total of 81 annual reports were available for review from the following grantees during the first and second rounds of document review: Action Health Incorporated (AHI), AfricMIL, Akin Fadeyi Foundation, Arewa Research & Development Project (ARDP), Association of Nigerian Electricity Distributors (ANED), Bayero University Kano, Brekete Family, BudgIT, Cable Newspaper Journalism Foundation, CEDDERT, CEWHIN, CHRICED, CITAD, Center for Transparency Advocacy, Centre for Socio-Legal Studies, CLEEN Foundation, Common Purpose, Community Life Partnership (CLP), Consumer Protection Council, CSAEFA, CSLAC, Daily Trust, HEDA Resource Centre, Human Development Initiatives, Imperial College London, International Centre for Investigative Reporting, International Federation for Women Lawyers, LANW, Moving Image, Nextier, NIALS, Partners West Africa, Pastoral Resolve, Premium Times Centre for Investigative Journalism, Public and Private Development Centre, Reboot, Right to Know, Sahara Reporters, Socio-Economic Rights and Accountability Project, Shehu Musa Yar'Adua Foundation, Stakeholder Democracy Network, Tiger Eye Foundation, Trust Africa, Wole Soyinka Centre for Investigative Journalism, WRAPA, and YIAGA.
2. *Grantee monitoring data collected in 2019 to address 12 measures of interest as identified through the grantee validation process*
 - Data deemed to be of usable quality were obtained for a total of 83 grantee monitoring indicators from 39 different grantees: AFCLSD, AfricMIL, Akin Fadeyi, ANED, Brekete, BudgIT, CEWHIN, CHRICED, CLP, CNJF, CODE, COGEN, CSACEFA, CSLAC, CTA, DTF, Equal Access, FOMWAN, GCC, HDI, HEDA, ICIR, LANW, LEDAP, LTLF, PARE, PPA, PTCIJ, PWAN, Scene One Productions, SDIC, SDN, SERAP, SMYF, Sahara Reporters, TEP, WOCON, WSCIJ, and YIAGA. Additional useful data from PPDC were identified during the November 2019 grantee convening.
3. *CLEEN Household Survey 2017 and 2018* (further detailed below)
4. *SDN Corruption Perception Index Report 2017 and Final Report 2019*
5. *Girl Child Concerns Baseline Assessment Report 2018*
6. Information on National HGSF Program rollout, directly provided by Imperial College-PCD in both 2018 and 2019

CLEEN Household Survey 2017 and 2018 Design and Sample: Criminal justice grantee CLEEN Foundation conducted surveys in 2017 and 2018, aimed in part at obtaining a better understanding and perception of citizens' assessment of the implementation of the ACJA. The survey included one set of respondents who were members of the general public, and another of members of criminal justice agencies, such as police, judges, prosecutors, and others. The 2017 survey included 4,489 members of the public and 610 criminal justice practitioners; in 2018, there were 4,539 public respondents and 618 practitioners surveyed. Enumerators conducted the survey in six states: Abuja Adamawa, Cross River, FCT, Imo, Kaduna, and Lagos. CLEEN researchers chose respondents through a multi-stage stratified random cluster sample.

CLEEN Household Survey 2017 and 2018 Content and Methods: Data collection consisted of in-home, face-to-face personal interviews. The questionnaire was administered in computer-assisted personal interviewing format, using tablet devices. CLEEN researchers assured quality through enumerator training, survey piloting, and spot-checking of 15 percent of each enumerator's interviews.

CLEEN Household Survey 2017 and 2018 Analysis: CLEEN Foundation created a report of findings after performing descriptive and inferential statistics to produce population-level estimates.

EnCompass team members conducted secondary data analysis to produce point estimates within a 95-percent confidence interval. Most data were disaggregated by state, and some were disaggregated by age, gender, education, and income. All disaggregations presented in this report were statistically significant. Analysis was conducted in Stata 14 software.

CLEEN Household Survey 2017 and 2018 Limitations: CLEEN Foundation used face-to-face personal interviews; conducting fieldwork in certain regions with sensitive security was challenging. Additionally, there are criminal justice practitioners who work in sectors that deal with sensitive information, and CLEEN enumerators had to obtain formal permission before interviewing some officials.



Corruption Indices

Content: The evaluation and learning framework specifies four corruption indices used to measure the extent of corruption in Nigeria based on: (1) the population's perceptions of corruption, (2) the degree to which the population has opportunities to participate in government processes, (3) the degree to which the population prioritizes addressing corruption, and (4) social norms surrounding corruption.

These indices are drawn from the following sources:

- World Bank Governance Indicators dataset's control of corruption indicator
- Gallup's annual survey of Nigerians' attitudes about the country's top priorities
- The Afrobarometer survey

- Open Budget survey

The selection of these indices reflects the evaluation and learning partner's recommendations based on the measures in the evaluation framework and needs the MacArthur Foundation articulated.

Methods: Afrobarometer and Gallup are citizen surveys producing population estimates representative of the Nigerian public, while the World Bank indicator is a composite indicator that uses multiple underlying data sources, including both representative and non-representative sources, which are rescaled to create the aggregate indicator. The Open Budget Survey assesses the degree to which the government provides opportunities for public participation in national budget processes. While, no updates to the indices were available at the time of this report, each of these data sources contains extensive methodological details available on its website.

Each source presents time series data with at least three data points. Because data are available for some indices reaching up to 10 years back, they provide an authentic, high-level understanding of the situation prior to the start of On Nigeria, both at baseline and the trends for the preceding years.

Limitations: Corruption indices produced by international organizations and publicly available provide an overall snapshot progress toward On Nigeria's ultimate goal of reducing corruption at the highest level. Reducing corruption at the country level is the result of a multitude of actors and actions, and cannot be attributed to—or indeed, achieved by—any single intervention. In the long term, at the point where trends in these indices corresponding to On Nigeria's period of activity are clear, the evaluation framework anticipates that contribution analysis may be able to help understand On Nigeria's unique role in changes in the level of corruption at this high level. Nonetheless, these indices must be interpreted with extreme caution because they reflect trends much broader than On Nigeria's sphere of control. They are generally produced for diagnostic, not program evaluation purposes. Best practice guidance within the field of corruption measurement and evaluation recommends that impact measures be directly linked to the reforms a program promotes;² as broad measures of the overall amount or perceptions of corruption in a country, these indices, by their nature, are not directly linked to any one set of anticorruption reforms On Nigeria supports. The extent to which corruption—an illicit behavior that, therefore, occurs out of the public sphere of directly observable actions—can even be measured is a topic that remains controversial among governance experts. Furthermore, as noted above, indices may not be produced by responsible organizations on a regular basis or made public on a timeline congruent with a program's timeline. Corruption indices, therefore, may not be sensitive to capturing programmatic successes and as such, could only be useful proxies for understanding long-term trends.