

ANNEX 18. EVALUATION DESIGN AND METHODOLOGY

DATA COLLECTION METHODS

In this evaluation, mixed methods were used to analyze the implementation of the ABC+ interventions. The quantitative method used an online survey, and qualitative methods included literature reviews, key informant interviews (KII), focus group discussions (FGD), Outcomes Harvesting (OH) workshops, and Most Significant Change Technique (MSCT). Using multiple data sources through mixed methods enabled the evaluation to cover a wide spectrum, conduct a more in-depth analysis, and achieve a more holistic understanding of data through triangulation (Yin, 2003). Some of these methods were conducted virtually, while others were conducted in person. The sequence of data collection was as follows: desk review, survey, KIIs, FGDs, OH workshops, MSCT, and evaluation of supplementary materials.

DESK REVIEW

The desk review of documents started early in data collection, and this involved reading reports, policies, plans, etc., to gather data to answer the evaluation questions. A template matrix to organize reviewed data was used to gather evidence from reading documents. Desk review data helped answer the evaluation questions across all IRs and informed the crafting of questions for KIIs, FGDs, and workshops for OH and MSC. See Annex 14 for the types of documents reviewed.

SURVEY

Simple surveys were administered online. Participants were teacher trainees, supervisors, school heads, LGU, and private sector focal persons. The survey tool gathered data related to the evaluation questions on the project's relevance, effectiveness, and sustainability. It included questions to measure project relevance, satisfaction, and effectiveness to recipients. There were also questions to measure the relevance of ABC+'s intervention to existing policies at the central office, regional, and division levels of DepEd.

KEY INFORMANT INTERVIEWS

KIIs were conducted via phone and virtual conferencing platforms. Featured questions were informed by desk review data and the project context presented by IP during the initial consultation of the evaluation team. Questions were emailed to informants should they be unavailable for calls or online conferencing.

FOCUS GROUP DISCUSSION

A total of ten FGDs were conducted online and on-site by gathering seven to eleven informants representing different stakeholder groups. Participants were grouped according to the topics they would discuss. One group comprised representatives of 10 cities (DepEd heads, trainers for instruction and materials development, community members (LGU representatives, etc.), eight (8) with school heads and teachers (i.e., 7 languages + Maguinanawn), and one FGD with BSU and WVSU. The FGDs talked mainly about the program's interventions and their perceived outcomes (see Annex 21).

OUTCOME HARVESTING WORKSHOP AND MOST SIGNIFICANT CHANGE WORKSHOPS

Outcome Harvesting Workshop (OHW) and Most Significant Change Technique (MSC) were conducted to complement each other. OH prompts were informed first by the desk review data, which was used for crafting workshop prompts, which key program organizers validated. Finalized versions were used in a workshop. Using meta cards, participants (social actors) representing different groups of intervention recipients were gathered in a virtual and in-person workshop to respond to the prompts and discuss their answers. Collected outcomes and outcome descriptions were verified and categorized (domains of change) by program implementers (change agents). Refer to Annex 16 for the verified outcomes.

The outcomes harvested and the activities that lead to them are appropriate springboards for the MSC stories of the participants. OH is good at identifying intended and unintended outcomes and gives a group perspective to outcomes or changes, while MSC presents a change through the perspective of an individual. While OH outcomes tend to be general, MSC data can be rich and vivid.

Picking up from the identified outcomes and domains of change in virtual and face-to-face OH workshops, participants in the MSC Level 1 workshop (i.e., teachers, school heads, community/Barangay leaders, and parents, etc.) were asked to write down their most significant change stories in relation to their experience with ABC+ activities or outputs. A template was used to help participants recall changes they experienced or witnessed in relation to ABC+ activities (see Annex 12). The template guided participants in providing details and structure to their stories. Participants were given a chance to write more than one story. Once stories were written, the group read their stories and chose the most significant story based on the prompt given to them. Annex 17 presents the 66 Level 1 stories from the regions and the selected 16 stories from the set.

The set of selected stories from Level 1 was passed on to Level 2 workshop participants (i.e., DepEd and ABC+ trainers, DepEd school heads and supervisors, and focal persons) who met online to verify the content of selected stories. After deliberating, they chose the six most significant stories from the set. Please see Annex 18 for Level 2 MSC Stories and the reasons for selecting them. These were passed on to Level 3 story selectors (i.e., Key leaders of DepEd and ABC+) who selected one most significant change story. In selecting MSC stories, Level 2 and 3 selectors were given a set of criteria they may use (see Annex 19). Emerging themes from the stories and their differences or similarities were generated to complement OH results and better understand the effects of ABC+ implementation.

EVALUATION OF SUPPLEMENTARY READING MATERIALS (SRM)

Initially, developed SRMs were supposed to undergo evaluation; however, upon learning from data collected at the outset that the SRMs had already undergone a long and stringent quality assurance as part of their development, it was decided to forego this data collection activity.

CLASS OBSERVATION

Originally, the plan was to conduct class observations to gather data on instruction and materials used. However, at the time of data collection, DepEd's moratorium for class observations was enforced, and schools were strongly advised to focus on learning recovery instruction and forgo other activities (refer to Annex 20, a sample of localized suspension of observations in Region V). At the same time, the pandemic was still ongoing, and class observations would have to be communicated in advance, with health protocols followed. These requisites could not be accommodated at the time of data collection.

SITE SELECTION CRITERIA

The main stakeholders were composed of the following: (a) DepEd central, regional, and local offices; (b) Local Government Units; (c) Parent Community Engagement (PCE); and (d) private sectors. The analysis was at three levels: individual, institution, and organization. The evaluation design included a selection of sites for KIIs and FGDs, and a list of participants (education leaders, teachers, school heads, parents, LGU, NGOs, and private partners) for the case studies for USAID/PH’s approval.

Specifically, the selection criteria were on two levels, geographic location with the associated linguistic focus and participant profile. For the first level of selection criteria for the regions, Bicol, Western Visayas, and Maguindanao Schools Division in BARMM, who were part of the project, were automatically selected. The next filter looked into the school divisions in the three regions that received the assistance of ABC+, specifically schools’ proximity to RO (ARPY2, p.9). A sampling frame of schools from the implementing partner guided the evaluation team for its specific selection criteria at the level of respondents, making sure the geographic locations and the seven languages used in the project were well represented. However, for KII, a census of high-level positions involved in early-grade literacy was interviewed (i.e., DepEd Central, Directors of Regions V, VI, and BARMM, chief educational supervisors, and educational supervisors in charge of early-grade literacy). Table 3 shows the summary of data collection methods and corresponding participants.

KII PARTICIPANTS AND COMPOSITION CRITERIA

KII requires in-depth information from informants who were selected based on their key positions (if from DepEd and ABC+) and intensive involvement in ABC+ interventions. See Annex 21 for the list of KII participants and their backgrounds.

FGD PARTICIPANTS AND COMPOSITION CRITERIA

FGD’s success depends largely on the varied backgrounds of informants and the perspectives they contribute to the discussion. Their response to one another’s answers also matters. To ensure that quality FGD is observed, the number of participants was limited from 6-12. A maximum of two representatives per stakeholder group were invited to participate. The FGD participants and their backgrounds are found in Annex 21.

Table 3. Summary of Data Collection Methods and Corresponding Participants

Method	Description
Document Review	Annual Reports, DepEd documents
Key Informant Interviews (KII)	Central Office, ROs, and DOs
Focus Group Discussions (FGD)	Seven (7) language groups composed of teachers, school heads LGU, and private sector focal persons
Simple surveys	Teacher trainees, supervisors, school heads, LGU, and private sector focal persons

SURVEY

Stratified random sampling was employed in selecting the respondents for the online surveys. Cochran's formula with a 95% confidence interval and a 3% margin of error was used in determining the sample size. Annex 3 presents the sampling procedures and template table showing the distribution of all participating schools, teachers, and desired sample size by geographical area. Sample selection was guided by proportional sampling to ensure that schools from Region V, Region VI, and selected schools in BARMM are fairly represented.

OH PARTICIPANT COMPOSITION AND SELECTION CRITERIA

OH participants were those closest to program implementation, such as field implementers (e.g., focal persons, teachers, trainers, writers and illustrators, QA evaluators, etc.). A few members who were knowledgeable about the program but detached from the intervention were also invited (e.g., program designers and key leaders).

MSC PARTICIPANT COMPOSITION AND SELECTION CRITERIA

These were informants that represented different groups in program implementation. The following are the project-level groups and the backgrounds of their members. MSC workshops, virtual and face-to-face, were held for each ABC+ beneficiary region.

Level 1: Intervention recipients and trainers: trained teachers and their school heads, parents of children who received SRMs, and community members representing a Division where OH was conducted.

Level 2: Intervention implementers from DepEd, ABC+, and partners: focal persons, regional or division trainers (different from the ones in Level 1), writers and illustrators, materials QA evaluators, etc., representing a Division where OH was conducted.

Level 3: Program Leadership from DepEd and ABC+: program designers, policy, and decision-makers, Central, Regional, or Division DepEd leaders, etc.

The evaluation team also collected additional information from school records, school improvement plans, learning recovery program plans, etc. Documents were sourced by taking photographs of documents with the permission of the school head. Electronic files were requested and sent to the team by email when possible.

The team also used Google forms and similar computer and mobile-aided measures, such as using a mobile phone and video conferences through Zoom and Google Meet.

The data collection phase began with four weeks of fieldwork, including two weeks of reviewing documents collected onsite. A visual representation of data collection and expected participants can be found in Annex 4.

DATA ANALYSES

QUANTITATIVE

The evaluation used descriptive statistics and analytic data visualizations for the quantitative data and disaggregated data for sex and location to the extent possible. The said approach to quantitative analysis was consistent with USAID's embedded monitoring and evaluation in the Program Cycle. Data from the survey were analyzed through descriptive statistics. Analysis of progress of ABC+ toward indicators from the quarterly/annual reports. However, the absence of needed data during the scheduled time for analyses precluded the comparison of division and school-level data at the start and midline and the determination of the effect size of interventions by cohort, divisions, sex, and location. Classroom observation was not also done due to the COVID-19 pandemic restrictions imposed by DepEd. Thus, employing descriptive statistics on classroom observation data collected during fieldwork was not included in the data analyses of this evaluation report. Other survey items are found in Annex 22.

QUALITATIVE

The qualitative data were recorded and transcribed in worksheets. It was subjected to content analysis using computer-assisted qualitative data analysis software (CAQDAS). The evaluation team coded the responses and computed the estimated reliability index. A constant comparison method complemented the index where the team checked the coherence of responses, codes, and themes to ensure that the results had no external threats to the audience of this performance evaluation. The team also used content analysis on the outcomes harvested. Qualitative data analysis was used for FGD, KII, MSC, and OH. Expected information was about themes gathered from the discussion to analyze the evaluation criteria indicators. In both quantitative and qualitative analyses, computer software such as the SPSS and NVivo was used for data processing and analysis. The codes used and reliability results are found in Annex 23.

TRIANGULATION

The evaluation team combined and compared the quantitative and qualitative analysis results. Comparing the two results provided a complete understanding of the ABC+ project regarding relevance, effectiveness, and sustainability.

EVALUATION DESIGN MATRIX

The design matrix in Table 4 summarizes the evaluation design and methods.

KNOWN LIMITATIONS TO THE EVALUATION

DESIGN

This performance evaluation has five potential limitations: reduced sample size, inaccessible sites, participant bias, ability to recall information for self-reported data, and limited access to documents. First, on sample size - some respondents of the KII, and FGD were not able to participate due to prior commitments, health reasons/COVID-19, weak internet connectivity, and power outages that resulted in a reduced number of actual participants compared with the number of invited ones. Second, some study sites were not accessible due to the distance from the ROs/DOs. For instance, areas visited in BARMM were limited to those in the city proper, Cotabato. Third, in the self-reported data, the survey, KII, and FGD relied on what respondents said. Responses depended on

what people could remember because these activities or events occurred in the past. Lastly, the access to documents where some project contact persons were probably no longer connected with their schools or transferred to other schools or offices (principal and other office staff assigned to another school) and needed more time to look for documents, which resulted in delays or limited access to documents.

DATA MANAGEMENT AND ETHICAL CONSIDERATIONS

The evaluation team ensured that ethical conventions in program evaluations were adhered to. The team secured informed consent from all participants in FGD and KII and all survey respondents as part of the data-gathering process. The evaluation team prepared the Participant Informed Consent Forms customized for various types of respondents. Moreover, the team observed cultural sensitivity and DNH principles in the entire spectrum of the evaluation process. The participants received information during the informed consent process about the: (1) purpose of the evaluation, (2) the extent of their involvement, (3) their right to refuse or discontinue participation, (4) the potential benefits of the evaluation, and (5) confidentiality of their responses. All interview data is stored on secure computers in Panagora's Manila office. The Panagora Home Office ensured that personally identifiable data were not included in any public reports, including the final reports submitted to the USAID DEC. As per the nature of this performance evaluation, interview data sets will not be submitted to USAID or any third party.

The following sections present the gathered data, analysis, and conclusions for ABC+Relevance, Effectiveness, and Sustainability. Results are organized according to the topline evaluation questions which are followed by data for the learning questions.