

A rocky road to evidence: Evaluating literacy programmes using a trust-based approach in a context of fragility


Authors:

Edoé D. Agbodjan¹
 Nassibou Bassongui^{1,2}
 Kablan P. Kacou^{1,3}
 Alexis Tabo⁴

Affiliations:

¹CESAG Business School,
 CESAG: Centre Africain
 d'Etudes Supérieures en
 Gestion, Dakar, Senegal

²University of Abomey-Calavi,
 Cotonou, Benin

³School of International
 Development and
 Globalization, Faculty of
 Social Sciences, University of
 Ottawa, Ottawa, Canada

⁴Center for Learning in
 Evaluation and Results
 for Francophone Africa
 (CLEAR FA), Dakar, Senegal

Corresponding author:

Edoé Agbodjan,
 edoe.agbodjan@cesag.edu.sn

Dates:

Received: 24 Apr. 2024

Accepted: 04 July 2024

Published: 25 Nov. 2024

How to cite this article:

Agbodjan, E.D., Bassongui, N., Kacou, K.P. & Tabo, A., 2024,
 'A rocky road to evidence:
 Evaluating literacy
 programmes using a
 trust-based approach in a
 context of fragility', *African
 Evaluation Journal*
 12(1), a758. <https://doi.org/10.4102/aej.v12i1.758>

Read online:


Scan this QR
 code with your
 smart phone or
 mobile device
 to read online.

Background: Evaluation in contexts of fragility and violence has recently received attention because of the increased complexity of conducting such evaluations. The use of digital tools has been advocated for conducting these evaluations, but with limited results.

Objectives: This article presents an in-depth analysis of combining digital tools with in-person activities to build trust and develop the type of human interaction required to improve the quality of evaluation design and implementation in the context of insecurity, fragility and violence.

Method: Data collection was conducted both offline and online. Enumerators collected data through face-to-face individual interviews, and statistical analysis was performed using STATA software version 17.

Results: The objectives of data collection were achieved at 99%, notwithstanding the challenging security environment. Several factors contributed to this achievement, notably our methodological framework based on trust building, digitisation and iterative programming. Despite this commendable performance, the overall efficiency was found to be 63%, indicating a potential for a 37% reduction in data collection time.

Conclusion: The proposed trust-based approach has been successfully tested to enhance the quality of baseline studies and establish conditions for the success of other phases of evaluations.

Contribution: This case study serves as an evidence to what we call a trust-based approach to the use of digital tools in evaluation processes. We contend that the effectiveness of digital tools in enhancing the quality of evaluation design, especially in the context of fragility and violence, hinges on their integration with face-to-face activities, trust-based human interaction and careful timing.

Keywords: evaluation; fragility and conflict context; digital tools; trust; in-person interaction; Burkina Faso.

Introduction

Burkina Faso is facing multifaceted security challenges that hinder the implementation of the Sustainable Development Goals (SDGs). Two thirds of the territory is severely affected by terrorism, and the country has experienced political turmoil because of several *coups d'état* ['stroke of state'] over the past 5 years. In this context of fragility and violence, building individuals' skills, especially that of young girls, and men, is paramount to saving lives and coping effectively with the uncertainty and unpredictability. For this purpose, the Netherland-based international non-governmental organisation (NGO), Aflatoun,¹ has implemented two programmes, namely the AFLATEEN+² and AflaYouth³ programmes, in the Sahel, North, North-Central and East regions of the country in partnership with local government and civil society organisations (CSOs). These programmes aim at improving the social and economic empowerment of adolescent girls and young women (aged 15–35 years), including internally displaced young persons. They do this by

1. About us – Aflatoun International – Child Social and Financial Education.

2. AFLATEEN+ is the first component of the programme targeting adolescent girls aged 15–18 years in post-primary and secondary education who face challenges in accessing educational services.

3. AflaYouth is the second component of the programme, targeting young women aged 18–35 years outside formal education system with the aim of improving income-generation abilities.

equipping them with the social, financial and gender skills to make informed decisions about their finance, education, future employment, and sexual and reproductive health.

This article analyses the mixed-method evaluation of these two programmes using a pre-post evaluation design. Overall, the evaluation of these programmes is carried out through a qualitative and quantitative approach. This article focusses mainly on the data collection process of the pre-post evaluation. The baseline (pre-post) studies were conducted throughout 2023.

Evaluation in context of fragility and violence has garnered greater attention over recent years (Hassnain, Kelly & Somma 2021). Indeed, settings of conflict and violence display complex dynamics. As such, they present several challenges to evaluation (Hassnain et al. 2021) and require contextual sensitivity and adaptability (Hassnain et al. 2021). And yet, evaluation designs remain predominantly static and mechanistic, assuming that the contexts are relatively stable (Hassnain et al. 2021). For instance, the use of technology constitutes a key area of progress in evaluating programmes in context of fragility and violence. However, reports on the use of technology in conducting evaluation in settings of conflict and violence focus predominantly on data collection. While digital tools are valuable input, interpersonal relations remain the main asset to successful evaluation endeavours (Hassnain et al. 2021). The interplay between the in-person interaction and the use of technology in improving evaluation has barely been analysed. This leads us to pose the question: how could digital tools be combined with in-person interactions to improve the quality of evaluation design in this context?

This article attempts to respond to this research question. It provides an in-depth analysis of the combination of digital tools with in-person activities to build trust and develop the type of interaction required to improve the quality of the evaluation design and implementation in the context of insecurity, fragility and violence, using Burkina Faso as a case study. Several digital tools were used in the framework of AFLATEEN+ and AflaYouth programmes evaluation. They include KoboCollect, Teams, WhatsApp, SharePoint, phone calls and data analysis software, namely STATA. Similarly, face-to-face activities were undertaken such as in-person training workshops, one-to-one meetings and in-person interviews. Alongside these, informal local information channels were continuously used. We analyse the effect of these combined activities in line with some key challenges entailed to the evaluation, such as the liaison with stakeholders, data collection, team safety management and safeguarding, and gathering lessons learned. We reflect on how the phasing of these types of activities helped improve the agility and responsiveness of the evaluation team in the face of unpredictable programme variables. We then argue that the power of digital tools in improving the quality of evaluation design in this context relies on the phasing of its use in combination with in-person activities.

Programme evaluation in Fragile context: The state of art

Whether in conflict-bound, fragile States or not, evaluation is primarily a contextual matter (Fitzpatrick 2012). Context shapes and influences evaluation approaches and design (Rog, Fitzpatrick & Conner 2012). This holds particularly true given that data collection and evaluation methods pose distinct challenges based on whether the point of context is conflict, fragile or violence. In what can be called 'normal setting' for evaluation, namely a 'peaceful country', there are several challenges. For instance, evaluation requires balancing costs against the usefulness of the results, which Patton (2015) refers to as 'utilization-focused evaluation'. In normal settings, evaluators ought to develop 'a working relationship with the intended users to help them determine the kind of evaluation they need' (Patton 2015:458). At this juncture, the context must inform the choice of evaluation method (Rog et al. 2012).

For example, many authors view randomised controlled trials (RCTs) as the 'gold standard' for evaluation (Banerjee & Duflo 2011; Duflo, Glennerster & Kremer 2006). Yet, there are several reasons why RCTs are not appropriate including ethical prohibitions, logistical impossibilities and some programmes such as environmental protection or national security (Newcomer, Hatry & Wholey 2015). More recent works on the decolonisation of evaluation, complexity and systemic approach call for more context sensitive, fit-for-purpose approach and openness to a diversity of epistemology, ontology and philosophy, in a quest of a strong objectivity rather than technical perfectionism (Chilisa & Bowman 2023; Mertens 2023; Parsons & Winters 2023). The latest approaches offer more room to integrate indigenous and marginalised population views into the evaluation design, data interpretation and result validation.

Moreover, recruiting and retaining participants is a crucial challenge for evaluations conducted in typical settings. The way programmes' beneficiaries are treated can directly affect their ability to consent, and to provide information during the evaluation process (Cook et al. 2015). Obtaining consent alone does not suffice for conducting high-quality evaluation, which could inform public policy (Patton 1997, 2015). Evidence shows that beneficiaries may provide their consent but later opt out or provide incomplete or incorrect information. That is why it is valuable to maintain the motivation of beneficiaries throughout evaluation. Cook et al. (2015) suggest using specific criteria to assess beneficiaries' motivation, including the impact of beneficiary literacy and comprehension, staff members' presentation and explanation of data collection tools, and the methods used for data collection and recording. Furthermore, evaluations that possess a higher level of precision, reliability and generalisability result in increased costs associated with time, finances and politics.

In addition to these evaluation challenges found in normal settings, those in fragile and violent contexts pose even more

difficulties. During times of conflict, the urgency imposed by the violence and humanitarian crisis often render the poor and vulnerable invisible, creating what we may coin an 'emergency paradox'. On the one hand, decision-makers require reliable data for informed decisions regarding the plight of the poorest and most vulnerable. On the other hand, data deprivation tends to be worse (Hoogeveen & Pape 2020). According to Hassnain et al. (2021), challenges in fragile and violent contexts are numerous, including identification and access to affected populations, methodological requirements, with particular attention to unintended effects, a lack of appropriate tools and resources and a lack of understanding of power dynamics. Data collection during conflicts is hindered by inadequate road quality, insufficient telecommunications infrastructure, and, in some instances, populations hostile towards representatives of the central government who provide limited essential public services. Collecting data in such situations not only poses logistical challenges but also residents in these regions may have little allegiance to these perceived 'hostile' government representatives (Hoogeveen & Pape 2020). Thus, in addition to the lack of motivation that one observes in normal setting, a dimension of allegiance further complicates data collection in a fragile and violent context.

And yet, there are no foolproof techniques for collecting data in fragile and violent context (Hoogeveen & Pape 2020). Like with evaluation in normal setting, the chosen methodology must be informed by the context (Rog 2012). Adaptation to the context does not imply simplification, it rather requires innovation. For instance, in contexts where security is a concern, time restrictions must be factored into the evaluation methodology (Hoogeveen & Pape 2020). Therefore, using a new questionnaire design with intelligent sampling techniques at the question level is necessary to address the challenge of administering a lengthy consumption module that takes several hours. In addition to the technical challenges of reducing the evaluation questions, fragile and violent contexts present challenges to social relations. Abma (2006) argues that the social relations between the evaluator and the beneficiaries influence the opportunities and limitations of evaluation practices. However, establishing a social relation between the evaluator and the beneficiaries of development programmes that could influence the quality of the evaluation is particularly challenging in conflict contexts because of security concerns.

Theoretical background

Drawing upon the state of art, we took a trust-based approach to ensure quality of the evaluation in Burkina Faso. Evaluation is a process of judging the quality and the results of an intervention to inform decision-making. Such a process of judgement involves various stakeholders mainly a commissioner, internal and/or external evaluators, users, the intervention beneficiaries and stakeholders, and decision makers. For instance, commissioners make decisions about the scope and outreach of the evaluation and its future utilisation. The evaluators are required to undertake the

evaluation by collecting information from all the stakeholders and analyse the data to make judgement, in line with the commissioners' indications (terms of reference). Thus, the evaluation process requires interaction between the various stakeholders and a confrontation of their views on the intervention. This interaction is supposed to be particularly strong in the case of participatory evaluation, such as the one undertaken in the framework of this mandate. Drawing upon the assumption that trust is fundamental to social interaction (Fukuyama 1995; Ito 2003), we hypothesise that trust is also a key ingredient for the quality of the evaluation process. Without trust, it will be extremely difficult to both access respondents and to gain the stakeholders' ownership over the evaluation findings.

It is widely agreed that in the context of fragility, violence and conflict, trust is a scarce commodity. In the field of psychology, Mayer, Davis and Schoorman (1995) define trust as a trustor's willingness to be vulnerable or to take a risk by giving some target (the 'trustee') control in a situation. Criticism, however, has been raised against the emphasis this definition places on the 'willingness to be vulnerable' as a component of trust, because trust as a social phenomenon has the potential to provide a trustor with the sentiment of safety, even though the outcome of the relationship can prove the trustor to be wrong (PytlakZillig et al. 2016). Social relationships are uncertain by nature, but trust is an ingredient that provides a form of tranquillity in the face of this uncertainty. Trust is thus an initial bet, a belief, on the trustee's ability (trustworthiness) to behave according to the expectations of the trustor. From a socio-economic perspective, trust is a variable, and not a state. It fluctuates from a negative dimension defined as defiance, to mistrust, to an extreme positive value which is faith (Servet 2006). Defiance, mistrust and faith refer to various degrees of trust that evolve in the individual and collective experience.

It is thus paramount to examine the determinants of trust and/or mistrust and how these factors could be leveraged for social interaction. Prior research by Mayer et al. (1995) has suggested that such trust can stem from the trustor's perceptions of the trustee's competence, benevolence, integrity, and shared identity and values, as well as from perceptions of contexts and structures (e.g. regulations) that would encourage the target to behave in a manner that makes the target worthy of trust. When these structures are not functioning, the veil of trust is torn. Other factors that appear to influence trusting behaviours include perceptions of the legitimacy of an institution or institutional actors, loyalty to another individual or institution that may have developed out of prior interactions and one's dispositional tendencies to trust. Consequently, research and theory have distinguished between interpersonal trust (trust between individuals), institutional trust (individuals' trust in various institutions) and interorganisational trust (trust between organisations or groups) (PytlakZillig et al. 2016). Each of these multi-level forms of trust might have differing bases (i.e. be based to various degrees on perceptions of competence or integrity, for example). Moreover, approaches to building trust are

driven by assumptions on the concept of uncertainty. For Ziegler (2003) trust is an attempt to reach the certainty that allows us to act, in an uncertain world. As such, when the concept of uncertainty is reduced to its probabilistic component, the interaction is viewed in terms of a contract with a varying level of computable risk. This level of risk determines the level of trust at stake. It also determines the type of mechanisms to put in place to build or maintain trust, such as the quality and quantity of information to be provided, incentives to put in place and collateral required, among others.

Methodology

The case study

AFLATEEN+ and AflaYouth programmes aim to enhance the social and economic empowerment of adolescent females and young women (aged 15–35 years), including young internally displaced persons. This empowerment is fostered by equipping them with crucial social, financial and gender-based competencies necessary to make informed decisions regarding finance, education, future employment, and sexual and reproductive health in Burkina Faso. The AFLATEEN+ and AflaYouth programmes are implemented by a network of NGOs, the Framework for Consultation of NGOs and Associations for Basic Education in Burkina Faso (CCEB), in collaboration with the Ministry of National Education, Literacy, and Promotion of National Languages, and the Ministry of Youth and Promotion of Youth Entrepreneurship. The programmes span the Sahel, North, Centre-North, and East regions of the country. They are implemented in conjunction with local authorities and civil society organisations.

Consultation of NGOs and Associations for Basic Education in Burkina Faso, specifically, is implementing a capacity-building programme for trainers by mainstreaming social, financial and entrepreneurship skills into the schools and vocational centres curricula. The objective of this comprehensive educational programme is to facilitate their personal growth and development. AFLATEEN+ aims to transform the lives of adolescents (girls and boys) aged 14–19 years through social and financial education and entrepreneurship with a gender perspective. It covers 40 post-primary and secondary schools, with a combined student population of 14 000. AflaYouth is additionally targeting the informal training sector in Burkina Faso, serving as an educational resource that aims to expand the income-generating abilities of 5700 young women. It provides access to training, support, mentoring and apprenticeships throughout their transition into the formal job market or when launching entrepreneurial programmes.

Evaluation design: Sampling

We employed a trust-based iterative sampling and data collection strategy to take into consideration the nature and special context of the evaluation, which is one of fragility,

violence and conflict (see Figure 1). The trust-based iterative sampling approach exhibits two main characteristics, specifically focussing on cultivating trust and the iterative programming to accommodate the inherent complexities of the environment under study. Firstly, trust-based iterative sampling aims upon fostering trust between the evaluator and all involved stakeholders. This trust encompasses interpersonal trust, institutional trust and trust in the dynamic nature of the surrounding environment. The establishment of such trust facilitates effective collaborative processes, minimises the withholding of information and mitigates declarative biases, which could otherwise compromise the quality of the sampling, the data collected and the acceptability of the evaluation's findings.

Secondly, trust-based iterative sampling implements a sampling methodology that undergoes continuous updates, guided by the evolving security concerns, persisting until the final day of data collection. It's worth observing that fragile, violent and conflict-affected contexts often have limited infrastructure, security concerns and highly volatile environments of distrust (Celestina 2018; Kroener, Barnard-Wills & Muraszkiewicz 2021) that make traditional sampling and data collection methods challenging. Hence, trust-based iterative sampling and data collection methods focus on flexibility, adaptability and continuous improvement. Traditional sampling methods often involve fixed and predetermined sampling designs, such as a predetermined and fixed population size, a predetermined sample size, a predetermined and fixed list of individuals to be included in the sample and predetermined and fixed methods of selecting individuals to be part of the sample such as random sampling, stratified sampling or systematic sampling. These methods are effective under the assumption of a stable environment or in contexts of laboratory experiments.

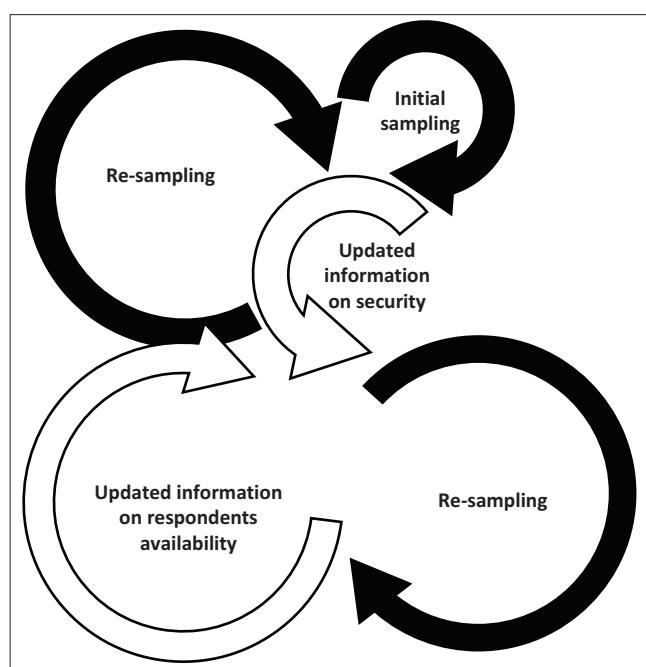


FIGURE 1: Iterative sampling model.

In context of fragility and conflict, however, a trust-based approach rather values a cyclical process of data collection, analysis and hypothesis refinement in which the evaluator continuously revisits and adjusts their methods and research questions based on emerging insights, allowing for the exploration of complex, dynamic, and evolving phenomena (Creswell 2003).

The reasons for choosing trust-based iterative sampling methods in this evaluation were numerous. Firstly, interactive sampling methods are characterised by greater flexibility and adaptability. They allow for the adjustment of sampling methods and strategies based on changing circumstances. This flexibility is crucial in areas where security, access and conditions can change rapidly. Secondly, iterative sampling methods request involving local communities. It may involve regular consultation with local communities to understand their concerns, perspectives and needs, ensuring that the sampling is more culturally sensitive and context specific. Thirdly, iterative sampling methods incorporate feedback and continuous learning. Sampling at each stage informs and refines subsequent sampling efforts. This feedback loop leads to improved sampling. Lastly, our sampling method relies on trust to mitigate potential biases that are likely to affect the quality of data such as withholding of information or information corruption. Beyond those key parameters of iterative sampling methods, we rely on the stratified sampling method by taking into consideration the spatial heterogeneity, gender diversity and status of internal displacement of the target population. Figure 1 presents the iterative sampling flow.

More specifically, our trust-based iterative sampling strategy followed 11 steps:

1. We initially aimed to cultivate a trusting environment with the primary stakeholders, notably the CCEB (*Concertation des ONG et Associations Actives en éducation de base*), Educo, and SUISSE SOLIDAR. To achieve this, we ensured the active participation of all local implementation partners in our online meetings, fostering their full engagement in the process. We also took measures to guarantee their comprehensive grasp of the evaluation's objectives, challenges and methodological approach by involving them in all inception workshops. We emphasised the importance of maintaining an updated record of potential and actual programme beneficiaries, considering changing environmental dynamics and unforeseen events. These concerted efforts were instrumental in building trust with the local implementation partners, who play a vital role in providing accurate and up-to-date administrative documentation, including the programme beneficiaries' list and information needed for population stratification, as well as security-related information across the implementation regions.
2. We requested the full list of the target population. The AflaYouth programme involved 358 participants while AFLATEEN+ targeted 2000 pupils and students.

3. We proportionally stratified the population according to four regions (North, Central-North, East and Sahel), gender (male and female) and status of internal displacement (internally displaced and non-internally displaced). Indeed, the proportional stratification followed the hierarchical structure of the population, that is 48 training centres⁴ are embodied in 16 communes, which are embodied in 11 provinces, which in turn are embodied in the 4 regions.
4. We determined a 'baseline' sample size using the above-mentioned information. That sample size was 404 for the AFLATEEN+ programme and 223 for the AflaYouth programme.⁵
5. We set an attrition rate to accommodate the changing context. We used a 20% attrition rate for AFLATEEN+ based on the internal displacement rate prevailing in the target regions. This information was received from local implementers. However, we set a 40% attrition rate for the AflaYouth programme based on the internal displacement rate recorded during the first phase of data collection for the AFLATEEN+ programme.
6. We proceeded with a random selection of participants.
7. We proceeded with a random selection of a replacement sample to accommodate the changing context. The size of our replacement sample was based on the attrition rate.
8. We communicated the list of participants to local implementers and data enumerators 2 weeks before data collection.
9. We updated the population size, the sample size and the list of participants after receiving updated information on the security situation from local implementers and data enumerators during the data collection training workshop.
10. We revised the population size, the sample size and the list of participants after receiving updated information on the security situation from local implementers and data enumerators at the end of the first day of data collection.
11. We updated the list of participants on a daily basis based on updated information on the security situation received from local implementers and data enumerators.

The progression in the sampling process is represented by the orientation of the arrows, where the flat ends lacking arrowheads signify the beginning of a task, and the arrowheads denote the end of a task, marked by the outcome. The arc on a white background denotes the initiation of the iterative sampling procedure, wherein the baseline information gathered is examined for conducting preliminary sampling, which will be continually refreshed and refined in the subsequent arcs. The white arrows indicate the acquisition of critical information within the sampling process, whereas the black arrows denote the outcome achieved subsequent to the incorporation of newfound vital information.

4. The 48 training centres included 39 centres for the AFLATEEN+ programme and 9 centres for the AflaYouth programme.

5. The details of the calculation of the optimal sample size are reported in Appendix 1.

Field preparation and negotiation

Setting up data collection period

There are no predetermined best periods for data collection in such a complex environment. Finding the right time to collect data requires some flexibility and constant monitoring of the situation in the field. As such, data collection for evaluations in fragile and violent contexts should ideally be planned for several months ahead of time, considering security risks and their dynamics in the study areas. Evaluators should constantly monitor information on security risks to guarantee the safety and well-being of both enumerators and respondents. Such a challenge is closely linked to the degree of trust established between the planner-evaluator and key evaluation stakeholders, which include local implementing partners, local authorities, schools and managers of professional training centres. In this context, once the trust of local programme implementation partners was secured, it became essential to earn the trust of the other parties. To achieve this, correspondence was sent to the local partners responsible for implementing the programme. The letter was drafted ensuring that the appropriate language and tone were used. This required proofreading by several various stakeholders to check appropriate wording and phrasing. The purpose of this letter was to inform the administrative authorities about the evaluation's timetable, participants, objectives and its goal. This initiative played a crucial role in fostering trust among administrative authorities, schools and professional training centres' managers, allowing them to take ownership of the challenges associated with data collection. Consequently, it facilitated the gathering of input from these parties for the adjustment of the data collection schedule in the light of security concerns and respondent availability.

Furthermore, it's worth noting that the complexity of the safety of evaluators relies on the fact that many factors influence security, notably the nature of the activities of the programme under evaluation, the general conflict situation in the field and the characteristics of the evaluators⁶ (Moss, Uluğ & Acar 2019). Hence, the best times for data collection are characterised by periods of relative calm or ceasefires, while considering the availability of respondents and potential exogenous climatic events such as rains, floods, high temperatures (Sahel region for example), and challenges relating to administrative procedure as well.

Concerning this evaluation, the initial timetable for the baseline data collection was planned for August 2022. However, this schedule was revised given the political and security challenges that prevailed in Burkina Faso. Given this situation, and after many meetings with the evaluation commissioner and other stakeholders, the evaluation timetable was revised, and the baseline data collection was rescheduled from 01 to 15 December 2022. Furthermore, because of many other challenges, the effective period of data collection was 27 February 2023 to 03 March 2023, indicating

6.These characteristics may include nationality, ethnicity, religion and gender of the evaluators.

a delay ranging from 5 to 8 months. There were various factors that contributed to the delay, notably the deterioration of the security situation in the Sahel regions leading to the closure of some schools, the time to hire local experienced interviewers, the time required to get administrative authorisations, the availability of respondents conditioned to the schools' calendars and internal administrative challenges (both and between the evaluator firm and the commissioner of the evaluation).

Preparing for data collection

Preparation activities for data collection are crucial to ensure accurate, relevant and reliable data. Whatever the type of data collection (surveys, interviews, observations), some key activities need to be carried out, notably the development of data collection tools (questionnaires, interviews guides, Focus Group Discussions [FGDs] guides, etc.), the internal and external revision of these tools, the recruitment and training of data collectors, and the pilot testing of data collection tools. Also necessary are information activities relating to local authorities, securing the necessary administrative, logistical and material resources (such as transportation, equipment, data storage and a budget) and the development of supervision and quality control plan. Our preparation activities do not differ *per se* from those implemented in traditional evaluations. However, the main difference was the way those activities were implemented. We focussed here on key activities such as the recruitment and training of enumerators, information activities relating to local authorities, securing the necessary administrative, logistical and material resources, and development of supervision and quality control plan.

Recruitment of enumerators and fieldwork facilitators

Similar to preceding phases, the establishment of trust within the diverse array of local stakeholders emerged as a pivotal factor contributing to the efficacy of the recruitment process for data enumerators and local facilitators. Notably, the trust gained through our interactions with local programme implementation partners and administrative authorities facilitated the delegation of the recruitment process to the local programme implementation partners. Subsequently, guided by the terms of reference and in collaboration with the local authorities, the local programme implementation partners adeptly conducted the recruitment of data enumerators and local facilitators. The facilitators were selected based on their social capital and ability to be heard by local authorities. For this reason, the majority of them are school teachers with good reputation.

The recruitment process for enumerators involved four stages:

1. submission of curricula vitae (CVs) of local enumerators by the PROMESSE programme implementation team and experienced evaluators in the region
2. examination of the CVs and selection of enumerators by evaluation experts

3. submission of the selected CVs to the PROMESSE programme implementation team for its assessment
4. online interview with those selected via the ZOOM platform.

These local enumerators represent 90% of the total number of enumerators. The remaining 10% of enumerators came from the evaluation team, notably research assistants from the CLEAR FA centre. This mixed composition of enumerators is intended to minimise the risks of displacement between regions while promoting the participatory and experience-sharing approach of the evaluation process with the local partners. One of the key innovations in our recruitment process was related to its reliance on an informal and local-based process rather than a formal recruitment one that would require a lengthy process in which local applicants may not be selected in the face of external, more competent applicants.

Despite the intricacies inherent to the evaluation environment, which presented challenges in the recruitment of local enumerators, we successfully formed a team of local, highly educated data enumerators with substantial years of experience. Moreover, our enumerator team demonstrates a balanced gender representation, with women comprising more than half of the enumerator team. Figure 2, Figure 3 and Figure 4 show the composition of enumerators by the level of education, gender and years of experience, respectively.

Following the compilation of the selected individuals, we organised two virtual meetings involving the data collection agents, local facilitators and the local programme implementation partners. The primary objective of these sessions was to reiterate the key aspects of the evaluation and data collection, as well as to provide a comprehensive contextual understanding of the evaluation process. Furthermore, these meetings highlighted the importance of sharing experiences, emphasising that evaluators could gain valuable insights from data enumerators rooted in their local experiences. Consequently, these multiple interactions yielded more comprehensive and up-to-date security information from the data enumerators concerning the programme's operational regions, prompting a revision of the data enumerators field appointments.

Additionally, the data enumerators proffered recommendations regarding the most effective and efficient modes of transportation for the fieldwork. These suggestions, in turn, contributed to the refinement of the transportation plan. Each local enumerator was affected in his region of residence to reduce the risks related to cross-regional displacements.

Training of enumerators

The training workshop was held on 25 February 2023 in hybrid mode, at a hotel in the capital city of Burkina Faso, Ouagadougou and online. This represented the inaugural

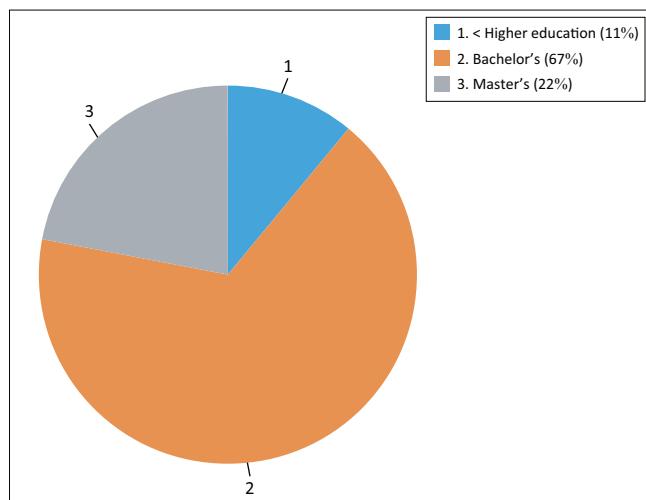


FIGURE 2: Enumerators' education level.

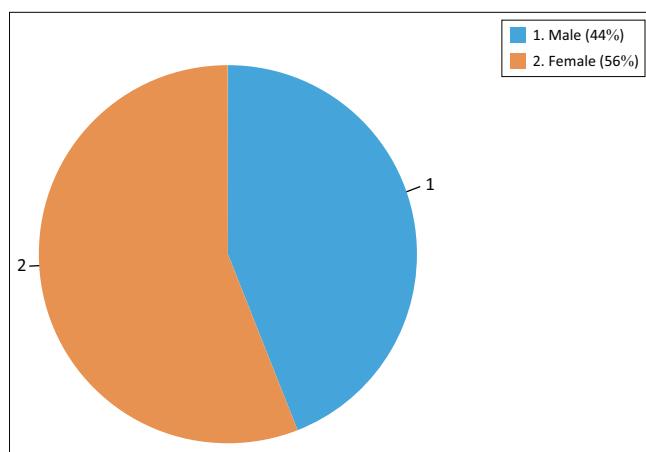


FIGURE 3: Enumerators' gender.

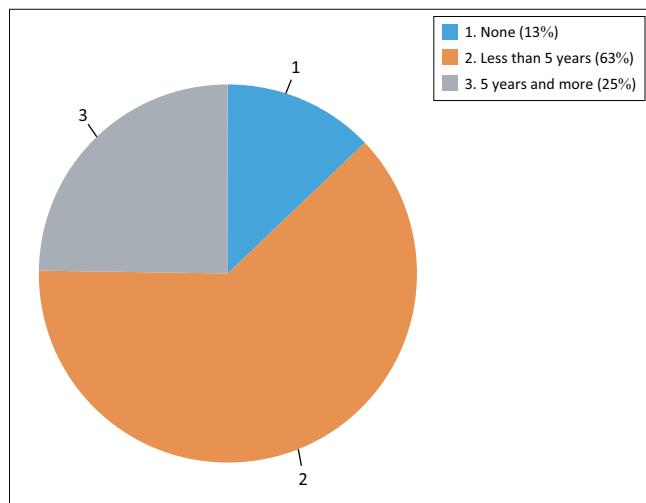


FIGURE 4: Years of experience.

face-to-face engagement conducted in the field as an integral facet of this evaluation. The workshop included local programme implementation partners, data collection agents and local facilitators. This interaction served as a platform for evaluators to programme a positive, non-virtual impression, thereby instilling confidence in these pivotal local stakeholders and reiterating their commitment to a candid

collaborative effort. To this end, the initial part of this workshop was dedicated to a 'getting to know you better' activity, followed by presentations from each of the participants. Furthermore, breaks, notably during coffee and lunch intervals, were strategically leveraged to foster informal conversations with all attendees. These informal exchanges effectively served to dismantle potential barriers that might impede the unrestricted expression of ideas and perspectives among the participants.

Overall, 20 participants attended the workshop, including 8 enumerators, 5 local facilitators, 1 focal point and 3 evaluators, including the lead evaluator. Two research assistants participated remotely. Local facilitators were selected based on their knowledge of the programme's activities and their respective regions of intervention. All facilitators were suggested by the local implementer partners. The content of the training workshop encompassed:

1. The opening ceremony that presented CCEB. This presentation was performed by CCEB focal point,
2. some fundamentals of the project, programme and public policy evaluation were presented to participants by the lead evaluator,
3. The objectives and methodology of the evaluation were presented,
4. In-depth presentation of the questionnaire and data collection digital platform. This session provided the enumerators and other stakeholders with the opportunity to explain their understanding of each individual question, and to rephrase them to better adapt them to the local realities. The participants were subjected to practical exercises through the administration of the questionnaire in local languages. The training also consisted of getting to grips with the digital data collection tool (KoboCollect) as the participants became familiar with the survey questionnaire,
5. The final session of the training workshop consisted of training and raising participants' knowledge and awareness in the ethical considerations governing data collection enumerators' responsibility for the quality of the data that will be collected. To this end, emphasis was placed on obtaining respondents' consent before every single interview. Similarly, interviewers were reminded that respondents have the right not to answer any question of their choice or to stop the interview at any time and that participants are responsible for guaranteeing good security conditions and protecting the integrity of respondents. Furthermore, information on the logistical arrangements required was shared with the participants.

Development and monitoring of a security map

The development of a security map was invaluable for this evaluation because the national security map was not accurate for our specific needs. Indeed, this map was aggregated, failing to provide detailed security information on our target areas. Besides, this nationally based security map was not updated daily. For these reasons, it was important to develop and update our own security map. We

TABLE 1: Disaggregated security map.

Region ID	Province ID	Commune ID	Training centre ID	Insecurity level
1	1	1	1	High risk
			2	
			3	
			4	
	2	2	5	Moderate risk
			6	
			7	
	3	4	8	High risk
			9	
			10	
2	5	7	11	
			12	
			13	
			14	Moderate risk
			15	
	6	9	16	
			17	
			18	High risk
			19	
			20	
3	7	12	21	
			22	Low risk
			23	
			24	
			25	
	8	15	26	Moderate risk
			27	
			28	
			29	
			30	
4	8	15	31	
			32	
			33	High risk
	16	16	34	
			35	
			36	
			37	
			38	
			39	

relied on the nationally based security map as a baseline map. Then, we used local informal informants to disaggregate the map down to our target areas (Table 1⁷). It can be seen from Table 1 that our map gives security information for training centres that were at the lowest level of spatial decomposition. This disaggregate map was updated continuously using our local informal channels.

Data collection tools and techniques

This phase represents the culminating stage of the data collection process, and its success is fundamentally contingent upon the trust established among local authorities, schools and professional training centres' managers, respondents, data collection agents and their immediate surroundings. To establish such trust, local facilitators, in collaboration with the appointed data enumerators in their respective region, engaged in physical meetings with various local authorities to

7.The names of regions, provinces, communes and training centres were removed to guarantee anonymity.

elucidate the objectives of the evaluation. Special emphasis was placed on clarifying that this evaluation remained entirely detached from political or governance concerns. Subsequently, prior to questionnaire administration, informal dialogues took place between respondents, local facilitators and collection agents to foster trust among the respondents. These preliminary exchanges prior to the questionnaire allowed respondents to ask questions related to the evaluation, including queries about the evaluators, before beginning the questionnaire. These preliminary informal conversations served to establish a foundation of trust with the respondents, thereby mitigating potential issues related to withholding of information or declaration biases. Furthermore, the establishment of respondent trust often extended into informal discussions occurring after the questionnaire introduction. Notably, data enumerators found that respondents frequently sought these post-introduction informal exchanges. During these discussions, our data enumerators were surprised by the degree of personal disclosure from respondents, particularly among internally displaced individuals. This phenomenon attested to the creation of conditions conducive to the survey as supported by scholars who highlight the paramount significance of trust in the research process, particularly in studies involving conflict-affected and marginalised populations (Hynes 2003; Lammers 2007).

This evaluation relied on a quantitative questionnaire eliciting responses to key evaluation questions. Given the complex nature of the evaluation environment, we digitalised the questionnaire through the KoboCollect platform. The digital questionnaire was then introduced face-to-face, using the offline digital tool. The reliance on local enumerators facilitated the access to respondents even in the face of relatively high insecurity contexts. The motivation for this choice was threefold.

Firstly, in contexts of fragility, violence and conflict, accurate and timely data collection is vital to reduce the impact of unexpected events (e.g. deterioration in the security situation) that may hinder the finalisation of the data collection. This is particularly critical in the context of this evaluation where al-Qaida and Islamic State-affiliated groups could attack the civilian population at any time. Secondly, our digital data collection tool enables real-time data collection, which is essential for monitoring rapidly changing situations and reducing the risks of a long stay in fieldwork for the safety of enumerators. Indeed, two digital data collection experts were assigned to supervise real-time data collection and all the submissions were checked in real time. Moreover, KoboCollect platforms can accommodate various data types, including text, images, audio, video and geospatial data, allowing for comprehensive data collection in complex environments. It allows for secure data storage, transmission and encryption, reducing the risk of data loss and unauthorised access. Besides, the possibility of collecting geospatial data was invaluable for mapping collection areas and tracking the effectiveness of enumerators' work. This

monitoring allowed us to ensure quality control and share the experience with all the enumerators. Lastly, our digital data collection tool offers offline data collection capabilities. This is invaluable in areas with unreliable Internet access, ensuring that data can be collected even when connectivity is limited or disrupted by conflict. That was very practicable in many regions, especially those in the Sahel and East regions. Apart from these features, our data collection relied on daily online reporting from each enumerator. The digital data collection experts who remotely monitored the data collection process synthesised the key challenges of each enumerator and a 1-h online meeting was taken to share those challenges, how they had been addressed and some propositions.

Communication and incentives

Effective communication and incentives have played a central role in fostering trust among the various stakeholders involved in the evaluation process, including the implementing organisation (CCEB), data collection agents and local facilitators. Despite the evaluators being international rather than national, there were no substantial barriers between the evaluation stakeholders and the evaluators. Consequently, we adopted a communication model designed to build trust, taking into account the diversity of the team. To achieve this objective, a range of communication channels and tools were employed, encompassing email correspondence, telephone conversations and online meetings conducted through platforms such as Zoom, Microsoft Teams, Google Meet, as well as communication via WhatsApp messages and calls. It's important to notice that the geographical locations of data collection agents and local facilitators posed challenges because of limited data connectivity as well as busy schedules, necessitating the adoption of this diverse array of communication methods.

In response to these constraints, targeted meetings were frequently organised for those unable to participate in the main meetings, utilising specific communication tools and designated time slots. For instance, individual telephone calls were conducted with data collection agents facing connectivity issues, security concerns or scheduling conflicts preventing their attendance at meetings. Targeted WhatsApp groups were created to maintain continuous communication. In addition, we ensured that a team member proficient in the primary local languages spoken by the data collection agents was part of the evaluation team to mitigate any communication issues arising from language barriers. Furthermore, other incentives, including per diem payments and transportation allowances, were disbursed directly to the data collection agents and local facilitators following the training workshops. This served to alleviate any financial constraints that might impede the fieldwork. Notably, the per diem payments were two-three times higher than the average rate in the absence of conflicts and insecurity. Furthermore, all data collection agents and local facilitators received their completion certificates without any complications after the data collection dissemination workshop.

Informed consent

As part of the *Centre Africain d'Etudes Supérieures en Gestion* (CESAG) ethical research principles, informed consent was obtained from all participants. Specifically, an informed consent form was directly submitted to respondents of the AflaYouth programme who were 18 years old or older. For those who were not literate, the interviewer read and explained the consent form in the local language before the respondent signed it. For the AFLATEEN+ programme, which involved adolescents under the age of 18, two informed consents were obtained. The first consent was from the adolescents' parents or legal representatives. This consent form was introduced by the local implementation partner of the programme, the CCEB. The second informed consent was obtained directly from the adolescents. Before administering the questionnaire, the interviewer read and explained the consent form to the adolescent, and the interview commenced once the adolescent gave their consent.

Ethical considerations

Ethical approval to conduct this study was obtained from the Centre Africain d'Etudes Supérieures en Gestion (reference no.: CESAG/DRI/01/2022).

Results

This section presents the main findings of the field data collection. Initially, we elucidate the quantitative goals set for data collection and the corresponding achievements. Subsequently, we offer insights into the gaps that emerged between the pre-defined objectives and the achieved outcomes, with particular emphasis on the contextual factors of the evaluation and the efficacy of our methodological approach.

Main findings of data collection

The objective for the data collection was to interview 404 participants from middle and high school student populations, distributed across the programme's four intervention regions. The first day of data collection was primarily dedicated to establishing connections between the data enumerators and administrative authorities within the education sector. With rare exceptions, most collection agents successfully established these connections within the first day. Consequently, after a span of 2 weeks dedicated to data collection, a total of 401 respondents were interviewed, thereby yielding an impressive completion rate of 99% (Table 2).

Equally, the pre-defined objective for the AFLAYOUTH programme component was to interview 206 girls and young women, positioned across the programme's four intervention regions. The initiation of data collection was primarily focussed on facilitating interactions between data collection agents and the management personnel of professional training centres. However, it is worth noting that certain data collection agents, particularly those in the North and East regions, successfully started on the second day of the data collection. The data collection process spanned 8 days, with region-specific achievements comprehensively outlined in Table 3.

TABLE 2: Data collection achievement for AFLATEEN+ programme.

Region	Achievement	Target	Performance (%)
North centre	80	83	96
East	73	73	100
North	181	181	100
Sahel	67	67	100
Total	401	404	99

TABLE 3: Data collection achievement for AflaYouth programme.

Region	Achievement	Target	Performance (%)
North centre	71	71	100
East	49	49	100
North	50	50	100
Sahel	36	36	100
Total	206	206	100

Achievement gaps

Overall, the data collection successfully met its predetermined objectives, notwithstanding the challenging security environment. Several factors contributed to the accomplishment of these objectives, notably our methodological framework grounded in trust building, digitalisation and iterative programming. Furthermore, online oversight and quality assurance meetings conducted with the data enumerators and local implementation partners played a pivotal role in enhancing the daily interview throughput. To illustrate, within the second phase of data collection, three online meetings were scheduled during the week of data collection, aimed at continually assessing the challenges faced by field agents and proposing remedial solutions in subsequent days. Despite our innovative methodological approach, the findings presented earlier demonstrate that not all data collection objectives were entirely fulfilled. In addition, even in cases where objectives were achieved, there exist avenues for improving the overall efficiency of the data collection activity.

Regarding the AFLATEEN+ programme, the quantitative data collection objectives were met at a rate of 99%. However, certain respondents could not be interviewed, primarily because of the insecurity context surrounding the data collection. Indeed, the high prevalence of internally displaced persons and the volatility of their movements posed challenges, notwithstanding the methodological provisions put in place. We adopted an iterative sampling approach to dynamically update the list of individuals to be interviewed in response to the evolving security landscape. Furthermore, our trust-based approach facilitated the active engagement of the local actors in verifying and confirming the up-to-date addresses of sampled individuals who had undergone residential relocations.

In contrast, for the AFLAYOUTH component, a 100% achievement rate was realised regarding the quantitative objectives for data collection. This exemplary performance can be attributed to several contributing factors. Firstly, it results from capitalising on the experiences acquired

during the first collection phase of the AFLATEEN+ component. It is essential to notice that the data collection for the AFLAYOUTH component was executed 3 months after that of the AFLATEEN+ component, affording evaluators, data collectors, local facilitators and other stakeholders the opportunity to assimilate and apply the lessons learned from the preceding phase.

Despite this relatively commendable performance, the need still exists to enhance the overall efficiency considering the achievements. Notable considerations include the planned duration of data collection for each of the two phases, where the first phase extended for 15 days in contrast to the second phase, which concluded within 8 days, resulting in efficiency of 33% and 63% for the initial and subsequent collection phases, respectively. The underlying factors contributing to the diminished efficiency of the data collection activity are multifaceted, encompassing administrative, logistical and technical challenges:

1. Data enumerators encountered difficulties in securing clearance from the administrative authorities within the education sector. Several agents reported that most administrative authorities in the education sector received notification of the data collection mission relatively late, thus hindering the timely launch of data collection.
2. Data enumerators also faced challenges in coordinating appointments with student respondents through school directors and principals. Many schools were concurrently immersed in examination activities during the designated data collection week, engendering a scheduling conflict between academic obligations and data collection efforts. This issue was explicitly documented in 14% of the daily field report⁸ submitted by data enumerators. The daily report form is presented in Appendix 2.
3. The arduous accessibility to schools, characterised by impassable roads, considerable distances between educational institutions, and the long detour routes for security considerations, presented a significant impediment to data collectors. This challenge was reiterated in 11% of the daily reports submitted by collection agents.
4. Additional impediments were attributed to the unavailability of local facilitators, whose role was pivotal in facilitating the introduction of collection agents to school principals and student members of the AFLATEEN+ club.
5. Furthermore, the high rates of school dropouts and internally displaced students, coupled with the closure of certain schools because of security concerns, led to more time invested in engaging with respondents. Despite the availability of replacement samples, daily reports statistics indicate that 59% of data enumerators acknowledged the need to replace one respondent each day, with an average of two replacements per day. Figure 5 highlights that dropout and absenteeism constitute the primary factors necessitating these replacements.

⁸A daily field report was prepared and submitted by each data enumerator using our digital tools, namely KoboCollect, WhatsApp, direct telephone calls. The choice of the tool depends on their availability at the time and place of reporting.

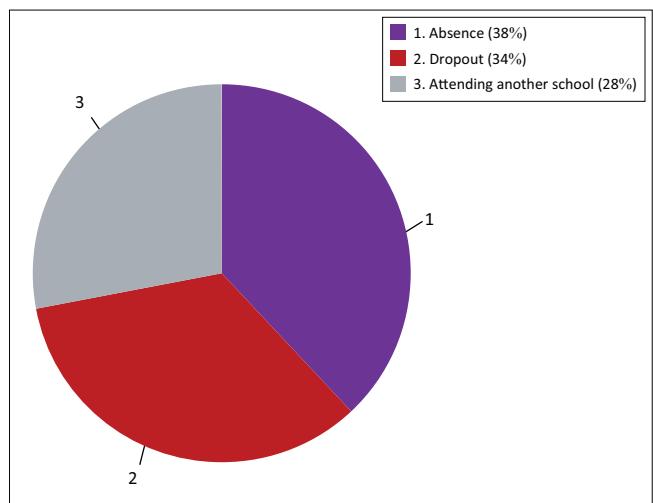


FIGURE 5: Patterns of replacement.

The enhanced efficiency observed in the second phase can be elucidated by several factors associated with the experience acquired from the first phase. Firstly, the scheduling of the timetable adopted a more participatory approach, wherein the active engagement of those overseeing professional training centres was solicited. This participatory dimension evolved as a product of earlier activities in cultivating trust in the evaluators' perspectives. Secondly, the process of selecting local facilitators was optimised through the selection of individuals with closer proximity to the intended target population.

Conclusion

This article analyses the mixed-method evaluation of the AFLATEEN+ and AflaYouth programmes using a pre-post evaluation design, carried out through a qualitative and quantitative approach. It reports mainly on the collection of baseline data. Given the situation of fragility, conflict and violence, our trust-based approach sought to create sustainable conditions for collaboration between the various stakeholders, for the success of the baseline study but also to establish the conditions for the success of the other phases to come. Several lessons can be drawn from this experience.

Both digital tools and human face-to-face interaction are crucial to establish a trusting environment between the evaluators and the diverse array of stakeholders, including local implementers, data collection agents, respondents and institutions. Technology and human face-to-face interaction are complementary. To leverage a hybrid strategy that combines the two, some key precautions are required, such as ensuring timely and comprehensive notification to both administrative authorities and respondents, engaging local stakeholders in the planning of data collection, fostering a sense of local ownership and collaboration. It also relies on simplifying and reformulating the evaluation inquiries in the most straightforward and concise manner possible to facilitate comprehensibility for data enumerators and respondents. Regardless of urgency or the level of experience

held by data collection agents, a pre-testing phase for the questionnaire is also paramount to gauge the sensitivity of the questions to the context. Similarly, special attention should be devoted to the training of data collection agents by allocating more time for the training sessions – 2–4 days may be reasonable. Moreover, ongoing reliance on local informants to source and provide up-to-date local security information.

Acknowledgements

The authors would like to acknowledge Miché Ouedraogo and Hadé Zougouri for their assistance in recruiting local data enumerators and facilitating access to relevant information, particularly in the challenging fieldwork environment.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

All authors contributed significantly to the creation of this article. E.D.A. contributed to the conceptualisation and methodology of the research, the analysis of the results, and the writing of the article and was involved in reviewing, editing, and supervision. N.B. and K.P.K. contributed to the conceptualisation, data curation, formal analysis, methodology and writing of the article. A.T. contributed significantly to data curation, formal analysis, methodology and writing.

Funding information

The authors disclosed the receipt of financial support from Aflatoun International for the research.

Data availability

The data that support the findings of this study are available on request from the corresponding author, E.D.A.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. The article does not necessarily reflect the official policy or position of any affiliated institution, funder, agency or that of the publisher. The authors are responsible for this article's results, findings and content.

References

Abma, T., 2006, 'The social relations of evaluation', in *The SAGE handbook of evaluation*, pp. 185–199, SAGE Publications Ltd, Los Angeles.

Banerjee, A.V. & Duflo, E., 2011, *Poor economics: A radical rethinking of the way to fight global poverty*, Public Affairs, New York, NY.

Celestina, M., 2018, 'Between trust and distrust in research with participants in conflict context', *International Journal of Social Research Methodology* 21(3), 373–383. <https://doi.org/10.1080/13645579.2018.1427603>

Chilisa, B. & Bowman, N., 2023, 'The why and how of the decolonization discourse', *Journal of Multidisciplinary Evaluation* 19(44), 2–9. <https://doi.org/10.56645/jmde.v19i44.919>

Cochran, W.G., 1963, *Sampling techniques*, 2nd edn., John Wiley and Sons, Inc., New York, NY.

Cook, S.C., Godiwala, S., Keeshawna, B.S., Powers, C.V. & John, P., 2015, 'Recruitment and retention of study participants', in K.E. Newcomer, H.P. Hatry & J.S. Wholey (eds.), *Handbook of practical program evaluation*, 4th edn., pp. 197–224. Jossey-Bass & Pfeiffer Imprints, Wiley, San Francisco, CA.

Creswell, J.W., 2003, *Research design: Qualitative, quantitative, and mixed methods approaches*, 2nd edn., Sage Publications, Thousand Oaks, CA.

Duflo, E., Glennerster, R. & Kremer, M., 2006, 'Using randomization in development economics research: A toolkit', *National Bureau of Economic Research* t0333, 3–87. <https://doi.org/10.3386/t0333>

Fitzpatrick, J.L., 2012, 'An introduction to context and its role in evaluation practice', *New Directions for Evaluation* 2012(135), 7–24. <https://doi.org/10.1002/ev.20024>

Fukuyama, F., 1995, *Trust: The social virtues and the creation of prosperity*, The Free Press, New York, NY.

Hassnain, H., Kelly, L. & Somma, S., 2021, *Evaluation in contexts of fragility, conflict and violence. Guidance from global evaluation practitioners*, International Development Evaluation Association (IDEAS), London.

Hoogeveen, J. & Pape, U., 2020, 'Fragility and innovations in data collection', in J. Hoogeveen & U. Pape (eds.), *Data collection in fragile states*, pp. 1–12, Springer International Publishing, Cham.

Hynes, T., 2003, 'The issue of "trust" or "mistrust" in research with refugees: Choices, caveats and considerations for researchers', *UNHCR Working paper No. 98*, pp. 1–25, UN High Commissioner for Refugees, viewed 29 December 2023, from <https://www.refworld.org/docid/4ff2ad742.html>.

Ito, S., 2003, 'Microfinance and social capital: Does social capital help create good practice?', *Development in Practice* 13(4), 322–332. <https://doi.org/10.1080/0961452032000112383>

Kroener, I., Barnard-Wills, D. & Muraszkiewicz, J., 2021, 'Agile ethics: An iterative and flexible approach to assessing ethical, legal and social issues in the agile development of crisis management information systems', *Ethics and Information Technology* 23, 7–18. <https://doi.org/10.1007/s10676-019-09501-6>

Lammers, E., 2007, 'Researching refugees: Preoccupations with power and questions of giving', *Refugee Survey Quarterly* 26(3), 72–81. <https://doi.org/10.1093/rsq/hdi0244>

Mayer, R.C., Davis, J.H. & Schoorman, F.D., 1995, 'An integrative model of organizational trust', *Academy of Management Review* 20(3), 709–734. <https://doi.org/10.2307/258792>

Mertens, D., 2023, 'The pursuit of social, economic, and environmental justice through evaluation: Learning from indigenous scholars and the fifth branch of the evaluation theory tree', *Journal of Multidisciplinary Evaluation* 19(44), 11–23. <https://doi.org/10.56645/jmde.v19i44.749>

Moss, S.M., Uluğ, Ö.M. & Acar, Y.G., 2019, 'Doing research in conflict contexts: Practical and ethical challenges for researchers when conducting fieldwork', *Peace and Conflict* 25(1), 86–99. <https://doi.org/10.1037/pac0000334>

Newcomer, K.E., Hatry, H.P. & Wholey, J.S., 2015, 'Planning and designing useful evaluations', in K.E. Newcomer, H.P. Hatry & J.S. Wholey (eds.), *Handbook of practical program evaluation*, 4th edn., pp. 7–35, Jossey-Bass & Pfeiffer Imprints, Wiley, San Francisco, CA.

Parsons, B.A. & Winters, K., 2023, 'Paradigm-based evaluation for eco-just systems transformation', *Journal of Multidisciplinary Evaluation* 19(44), 24–44. <https://doi.org/10.56645/jmde.v19i44.799>

Patton, M.Q., 1997, *Utilization-focussed evaluation: The new century text*, 3rd edn., Sage Publications, Thousand Oaks, CA.

Patton, M.Q., 2015, 'The sociological roots of utilization-focussed evaluation', *The American Sociologist* 46(4), 457–462. <https://doi.org/10.1007/s12108-015-9275-8>

PytlikZillig, L.M., Hamm, J.A., Shockley, E., Herian, M.N., Neal, T.M.S., Kimbrough, C.D. et al., 2016, 'The dimensionality of trust-relevant construct in four institutional domains: Results from confirmatory factor analyses', *Journal of Trust Research* 6(2), 111–150. <https://doi.org/10.1080/21515581.2016.1151359>

Rog, D.J., 2012, 'When background becomes foreground: Toward context-sensitive evaluation practice', *New Directions for Evaluation* 2012(135), 25–40. <https://doi.org/10.1002/ev.20025>

Rog, D.J., Fitzpatrick, J.L. & Conner, R.F., 2012, 'Context: A framework for its influence on evaluation practice (Editor's notes)', *New Directions for Evaluation* 135, 1–6. <https://doi.org/10.1002/ev.20023>

Servet, J.-M., 2006, 'Banquiers aux pieds nus', *La microfinance [Barefoot bankers. Microfinance]*, Odile Jacob, Paris.

Ziegler, A., 2003, 'Susciter la confiance', *Foi et Economie foi et économie*, Paris.

Appendices starts on the next page →

Appendix 1: Sampling strategy

Sample size

We used the random sampling size calculation formula for a finite population (Cochran 1963):

$$\text{Sample size} = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \frac{z^2 \times p(1-p)}{e^2 N}}, \quad [\text{Eqn 1-A1}]$$

where:

N : denotes the population size (358⁹ for the AFLATEEN+ programme and 2000 for the AflaYouth programme);

e : represents the margin of error (5%);

z : score corresponding to the standard deviation of a given proportion from the mean (1.96, for a confidence level of 95%).

Furthermore, we anticipated a 20% attrition rate because of the volatile security context. This rate also makes it possible to compensate for experimental mortality during ex-post collection while maintaining a representative sample. Application of formula (1-A1) leads to a sample size of 223 individuals.

Sampling fraction

The sampling fraction is the proportion of the population selected for inclusion in the sample. It was determined by the ratio between the sample size obtained from equation (1-A1) and the total number of beneficiaries of each program. The sampling fraction for each stratum was determined using the formula:

$$\text{Strata sample fraction} = \frac{n_s}{N} \times n, \quad [\text{Eqn 2-A1}]$$

where:

n_s : represents the size of the stratum s , for all $s = 1, \dots, S$;

n : denotes the sample size determined previously in (1-A1).

We defined two levels of strata: regions and schools.

9.The AFLATEEN+ programme initially had a population size of 400. However, a few days before the data collection, our field focal points informed us that some training centres in the SAHEL region were compromised. Considering this information, we revised the target population from 400 to 358.

Appendix 2: Daily field report form submitted by data enumerators

30/01/2024 09:44

RAPPORT JOURNALIER

RAPPORT JOURNALIER

1. Date

yyyy-mm-dd

2. Code de l'enquêteur

3. Région assignée à l'enquêteur

- NORD
- CENTRE-NORD
- EST
- SAHEL

4. Nombre d'enquêtés enregistré

5. Avez-vous remplacés un/des enquêté(s)?

- Oui
- Non

6. Si oui, combien?

7. Cause du remplacement

8. Difficultés rencontrés