In the book Evaluation Failures edited by Kylie Hutchinson, 22 esteemed evaluators share funny, brave and highly insightful accounts of the lessons they have learned by making mistakes. The concept is liberating but is it one that African evaluators can easily adopt in a context and culture where we have been conditioned to hide our mistakes.
Introduction

The current open debate about failure in Monitoring & Evaluation stems from the book Evaluation Failures, edited by Kylie Hutchinson, in which seasoned evaluators reflect on some of the mistakes they have made and what they have learned from them.

The book originated from the 2015 American Evaluation Association’s conference themed Exemplary Evaluations in a Multi-Cultural World: Learning from Evaluation Successes. Noting the event’s evident bias towards documenting success, Michael Quinn Patton proposed a session on failure – in order to ‘bring some balance’ to the theme, as he writes in the book’s introduction.

The idea of balance is fundamental in moving discussions on Made in Africa evaluation forward. Of the 22 cases from around the world that were featured in the book, only two were from African evaluators.

The reasons for evaluation failure in the Global South, specifically in Africa, are vastly different from the rationale in the Global North. Yet, African stories of evaluation failure remain largely untold and undocumented.

This emanates in part from the continent’s long tradition to keep mistakes hidden – a direct result of the imbalance. If donors keep giving precedence to Western evaluators or evaluation consultancies over African evaluators when commissioning large-scale evaluations, it would be counterintuitive to talk openly about where they went wrong. Yet, this is not helpful. It is an iterative cycle that will very likely prevent the continent’s evaluators from learning from evaluations that went awry.

To encourage the sharing of stories, Jennifer Bisgard from African M&E firm Khulisa Management Services, chaired a panel at the 9th AfreA Conference on Evaluation Failures Made in Africa, collaborating with three senior African evaluators, namely Felix Muramutsa, Zakariaou Njoumemi, and Ahmed Ag Aboubacrine.

This article highlights the many unique evaluation challenges that the continent faces and shares four stories of failure by African evaluators to reduce the sting from failures that were ‘Made in Africa’ and encourage the confidence to voice them.

Key Messages

- Borrowing directly from the Evaluation Failures approach, this article shares four stories of failure by African evaluators (two of these are excerpts from the book) to reduce the sting from failures that were ‘Made in Africa’ and encourage the confidence to voice them.

- Similar to the book, it derives the lessons learned from making mistakes in four different regions of the continent – Southern Africa, East Africa, West Africa and the Sahel.

- The article concludes that the inability to value indigenous knowledge, contributes to typical Made in Africa evaluation failures, and therefore encourages African evaluators to admit to, and talk openly about failures for improved development effectiveness.
African continent is facing, including data transparency, the relationship with government stakeholders, resource constraints and the need to understand cultural contexts. It adopts a similar approach used in the Evaluation Failures book—presenting case studies of how evaluations went wrong in four African regions (Southern Africa, East Africa, West Africa and the Sahel) with a view to documenting the stories of African-rooted evaluation failure.

Evaluation failure: Southern Africa

Data sensitivity in corporate entities in South Africa

Across Africa, evaluators and program managers assume that data is being collected routinely and effectively. Is this assumption true? If it is not true, why do we hide it? Is there any crisis of confidence? Are there remedial actions? In this case study from South Africa, Jennifer Bisgard shares some of the learnings around failing to obtain evaluation data.

Both the evaluand (the NGO we were evaluating) and the donor assumed that, in our capacity of evaluators with the required technical skills and knowledge, we would be able to produce outstanding outcomes from the evaluation process.

Perhaps we could have reconstructed this data, but we were hit by another false assumption. Accustomed as we were to evaluating donor projects, we were completely unprepared for a key stakeholder group to refuse to cooperate and share their data with us.

The project was designed to provide opportunities for small businesses to be certified and thus become suppliers to large corporate entities from 2012 to 2015. These firms later joined the NGO as members and sat on their board. The NGO aimed to create “matches” between their 27 corporate members and approximately 400 certified suppliers.

At the donor inception meeting, they stressed that they needed quantitative data for financial results. One reason for the evaluation was that the NGO’s routine monitoring had reported low numbers of “matches” between corporate members and certified suppliers.

Thus, the donor required the evaluator to request corporate procurement divisions to share procurement data for the 2012-2015 period (including overall procurement expenditure, and certified supplier expenditure contrasted with non-certified supplier expenditure).

Upon request by the NGO, few actually provided the information. However, both the donor and the NGO assumed that corporate members would share this data more openly with an external evaluator, even though they had been unwilling to share it with the NGO itself. This assumption was not just false – it was incredibly false.

We spent weeks struggling to set up key informant interviews with corporate members, often running into a brick wall. When we finally got around to analyzing the routine data we, realized, to our absolute horror, that the data was incomplete, with most members only reporting erratically. The NGO also had a “data crash” – the firewall, antivirus, and back-up had failed and the repair of the historic dataset was grossly incomplete. Not only was there an abysmal lack of data, but also the little information that existed was lost in the crash.

Therefore, relying on the corporate members’ cooperation to share historic procurement data became paramount. However, we continually ran into confidentiality concerns,
indifference, and corporate staff who had never heard of the NGO. Many corporate members were unwilling to divulge the requested information, as they considered it too sensitive and confidential to share for an evaluation report that would end up in the public domain.

Things went wrong because the assumption that corporate entities would share their procurement and finance data was central to the verification portion of the evaluation, leaving us without an approach to attribute (or even contribute) towards the NGO’s impact. Although the corporate members themselves governed the NGO, even the board could not convince its members to release the information to our evaluation team.

We overestimated the corporate members’ commitment to the NGO and underestimated the sensitivity of the procurement and financial information. Although we signed confidentiality agreements, and cajoled and applied pressure, none of it was enough.

In the end, we succeeded in getting only partial datasets from six corporate members. We used this limited procurement data to triangulate with the spotty routine reporting data. We were able to answer most evaluation questions, but there was not sufficient data to document the impact.

A critical lesson stemming from this evaluation failure is that an evaluability assessment is essential prior to commissioning any evaluation. If, in this case, the donor had conducted an evaluability assessment, there would have been an early realization that an external evaluation could not address the extreme data gap. In other contexts, evaluability assessments can identify when the program has not progressed enough to be evaluated.

* This is an edited version of one of the case studies that appeared in Evaluation Failures: 22 Tales of Mistakes Made and Lessons Learned

### Evaluation failure: East Africa

#### Navigating conflicting stakeholder requirements in Rwanda

During the collection of baseline findings in an African context, the relationship between donors and governments can be particularly complex. Felix Muramutsa shares how he attempted this challenge with an evaluation in Rwanda.

The evaluation was a baseline prevalence survey of a child labor project run by an international NGO (INGO) for which I was working as Senior Monitoring & Evaluation Advisor and Deputy Director. The purpose of the evaluation was to assess the prevalence of child labor in the project’s intervention area. The major stakeholders were the donor, the Government of Rwanda (represented by the Ministry of Labor), tea companies, tea cooperatives, district and local authorities, the implementing local NGOs, the consultancy company, parents and children.

My role was to coordinate the overall baseline survey and ensure the consultancy company hired to conduct the survey delivers the goods accordingly, while fully taking into account the expectations of the donor, INGO, and the government. The role of the public staff was to ensure the baseline data reflects the local reality. As the overall baseline coordinator, I was expected to ensure the harmonization of all stakeholders’ roles and expectations.

As it is often the case in African evaluations, the requirements set forth by both the donor and INGO were different from those by the government; and as an evaluator, my job was to respect them all.
For example, the donor considered that the baseline results were first of all their business only. However, the government also required me to inform them of all the steps occurring and expected me to validate the findings with them prior to publication. I was instructed by my employer (the INGO) to not share the preliminary findings with the government, which were instead discussed and validated between the consultancy company, INGO and the donor. Later on, INGO shared a final version with the government, and I was tasked with following up with dissemination.

The alert came when project stakeholders started challenging our results during meetings with government officials. To my dismay, the government questioned the methodology, sampling and findings, and ultimately did not validate the report. I was personally caught in the middle, with each side requiring me to convince the other of their position.

I believe things went wrong when I first accepted instructions from the donor not to involve the government in all steps of the baseline survey. I should have known there would be complications at the validation stage. What contributed to this mistake was my blind respect for the donor and INGO's request to be excluded in the first place, and then include the Government, all the while knowing it conflicted with its requirements of being involved at all stages of the project, including the baseline step.

When things began to go wrong, I became a scapegoat. I was very frustrated to be treated by some government officials like an accomplice of “foreigners who want to spoil the country’s good image by exaggerating a fake child labor prevalence.”

To address the situation, I personally engaged, both formally and informally, government officials and local project implementers in close discussions and some advocacy to better understand the issues at play beyond technical ones. We discussed what improvements were needed for the draft report.

As an evaluator, my work was significantly influenced by the situation. It has sharpened me to better understand that the early involvement and active participation of all stakeholders is key for a successful evaluation.

For multi-stakeholder evaluations of this nature, it is also necessary to set up a joint baseline steering committee. The role of the committee would be to anticipate and discuss any technical, administrative or political issues that might surface at all levels, and then build the foundation of the next steps during research.

* This is an edited version of one of the case studies that appeared in Evaluation Failures: 22 Tales of Mistakes Made and Lessons Learned

**Evaluation failure: Central Africa**

**Failing to use mixed methods in a water project in northern Cameroon**

Resource constraints can hugely detract from the quality of evaluation findings. This case study from Northern Cameroon by Zakariaou Njoumemi details why it is critical to factor in both quantitative and qualitative methodologies with long-term social sector projects in Africa. Failing to do so will lead to a wrong response to why a project did or did not work.

In many rural communities in Northern Cameroon, access to safe drinking water is severely challenging. Many women and girls are used to walk long distances, often 10 to 15km away from their houses, to collect unsafe drinking water in rivers.

In the late 1990s, a five-year drinking water supply project was implemented by an
international NGO whose mission was to build several drinking water sources at central areas in different rural communities in Cameroon. The objective was to provide safe drinking water to every household, thereby reducing distances and travel times to drinking water sources.

By early 2000, a final evaluation was commissioned. The implementers were excited to know the proportion of women and girls who were using the newly created drinking water delivery points and their level of satisfaction.

We were appointed as junior evaluator to conduct a final evaluation of the project. We only used a quantitative approach, because the project management team argued that they had a limited budget for evaluation. I failed to convince the project management team to increase the evaluation budget in order to allow the use of mixed or combined quantitative and qualitative evaluation methods, even though I knew from experience it was critical for this evaluation.

Our evaluation design and methods therefore excluded a mixed methods approach. The evaluation findings that were released and disseminated to the implementers indicated that about 49% of women and girls were using the newly built drinking water sources located within the communities, while 51% bypassed them and continued to walk long distances in search of drinking water.

The quantitative methods could not shed light on why women and girls did not use the newly built drinking water delivery points situated around or near their houses. Based on these challenges, a qualitative evaluation had to be commissioned to complement the quantitative findings.

The qualitative evaluation discovered that in the assessed rural communities, women and girls were often confined to their houses within closed fences, which led to reduced access to recreational activities. It is only when they went out to collect drinking water that they had the opportunity to socialize, meet their friends and other members of the community, or attend informal meetings dedicated to women and girls. Therefore, walking long distances to rivers provided them with many social advantages.

Our failing to conduct a comprehensive evaluation taught us some lessons, the key one being that social sector evaluations in Africa should investigate much further than just what happened during project implementation, focusing above all on why and how it happened. The evaluator should always use a mixed methods approach or a combined quantitative and qualitative design to determine this.

Managers of development projects need to collect regular routine data through monitoring activities, while ensuring an adequate budgeting for evaluations, so as to cover both quantitative and qualitative evaluation designs and methods from the onset.

Meta-analysis of evaluation failure: The Sahel

Evaluation will not succeed without bearing in mind local context and indigenous knowledge

Ahmed Ag Aboubacrine refers to evaluation failures that can occur if evaluators are not able to meet with all the intended beneficiaries of a project or observe the actual changes due to seasonal factors such as migration or rain. He further highlights three contributing factors for why evaluations fail in the Global South, alluding to typical Made in Africa lessons that could be derived from these factors.
In many African countries, farmers usually engage in income-generating activities other than agriculture after harvest. This includes migrating to neighboring regions or countries to work as manpower. In some countries such as Guinea, Mali, Mauritania and Niger, which are occasionally stricken by drought and where farms are sometimes infected by worms, locusts or other plant diseases, the migration of labor force happens earlier in the farming season. That is why evaluators must go to the field to better understand such context-specific and exogenous factors that may bias their findings and conclusions.

What is the specific Made in Africa evaluation lesson that could be learnt? Asymmetry of information and knowledge undermines the likelihood of an evaluation’s success and usefulness. One major lesson is that in Africa, the usual evaluation planning is not enough to detect the real issues and lessons. One needs to know the context beforehand and include it in the evaluation design. To be successful, evaluations in Africa ought to include a preliminary field phase for preparation purpose.

Evaluators should not rely on desk work or completion reports from executing agencies to plan their evaluation exercise. They should go to the field and understand the social dynamics and seasonal effects before deciding how to conduct evaluations.

Failed evaluations in the Global South are attributed to several factors, one of the most notable being the use of Western biases by evaluators (from the West). Even if foreign evaluators do have the adequate expertise and experience, they often tend to look for the change patterns expected by the design (strategy, theory of change, results framework) of their respective organizations. Such a misconception is particularly pervasive in the United Nations’ common system and in bilateral agencies.

A second factor is the negative use of innovation in development interventions. With the emergence of results-based financing, there is a risk that agencies will create incentives for ‘cooked’ results. Implementing agencies who want money would focus on creating outputs for the sake of payments and therefore undermine the real relevance of their operations (whether these were the right ones or not).

A last enabling factor of evaluation failure the inability to value indigenous knowledge, and this is more relevant when it comes to typical Made in Africa evaluations. Many evaluations in the Global South miss the assessment of changes in local dynamics and power relations. The undocumented nature of these dynamics and relations makes it difficult for evaluators who have not been exposed before to the context to grasp the true nature of the changes that occurred.

In conclusion, as African evaluators or evaluators working in Africa, we need to take the sting out of the word “failure” and talk openly about and invest in understanding the specific circumstances that make evaluations unsuccessful, so as to set us on a path of learning, improved implementation and development effectiveness.
References


Nicola Theunissen is an independent communication specialist, with a focus on Monitoring & Evaluation. She has consulted for the African Evaluation Association (AfRE), Khulisa Management Services, and the South African Monitoring & Evaluation Association (SAMEA), amongst others. She also conducts work in the media industry, including for the fact-checking organization Africa Check. Nicola holds a master’s degree in Communication and an honors’ in Community Psychology.

Jennifer Bisgard co-founded Khulisa Management Services in 1993. As an expert in monitoring and evaluation (M&E) and organizational development, She leads evaluations and capacity building assignments in the education, and democracy and governance sectors. Prior to establishing Khulisa, Jennifer was a Senior Education Specialist at USAID/Pretoria from 1988 to 1993. She has served as board member for the African Evaluation Association (AfRE), the International Organisation for Cooperation in Evaluation (IOCE) and SAMEA. She holds a master’s degree in Social Change and Development from Johns Hopkins University.

Felix Muramutsa is a freelance consultant with 23 years of experience acquired mainly from more than 30 assignments in research, monitoring and evaluation. He has worked with UN agencies, international and local NGOs, civil society organizations, government institutions, foreign and national universities, and the private sector. Felix holds a master’s in Psychology and a Bachelor of Arts in Education.
Zakariaou Njoumemi is a senior health economist specialized in health financing and monitoring & evaluation of public health programs. He holds a PhD in Public Health, with specialization in Health Economics, from the School of Public Health and Family Medicine, University of Cape Town, South Africa. He is a senior lecturer/researcher in public health evaluation, health economics, health policy and health systems at the Department of Public Health, Faculty of Medicine and Biomedical Sciences at the University of Yaoundé. He is also the Director of Health Economics & Policy Research and Evaluation for Development Results Group (HEREG) in Yaoundé, Cameroon.

Ahmed Aboubacrine is a seasoned development evaluator with 17 years of experience, including for CARE International in Mali, Burundi, Sierra Leone, and Liberia, where he played a critical role in evaluation capacity building, program quality and impact measurement. In July 2010, he joined the Islamic Development Bank’s evaluation department based in Jeddah. Ahmed holds an Associate Degree in Mathematics, a bachelor’s degree in Applied Statistics and a master’s degree in Decision-Making. His areas of expertise include the design, monitoring and evaluation of relief and development interventions in the public and private sectors.