



CASE STUDY

Trusting the Data We Collect: Why Independent Quality Assurance Matters

A Professional Field Reflection

The author has extensive experience in humanitarian research, quality assurance, and Monitoring, Evaluation, Accountability and Learning (MEAL) across Somalia and the wider East Africa region, and is the founder of RAYID.

Introduction

In the contemporary humanitarian context, data is frequently described as essential foundation for situational awareness, coordination, and crisis management. Evidence is defined as any form of knowledge of sufficient quality to inform a choice and it has become the primary driver for guiding life-saving actions. Data matters because it provides the basis for determining the severity of a crisis, identifying the most vulnerable population groups, and documenting human rights violations. Without this information, the humanitarian community risks operating in a vacuum, relying on assumptions that may fail to reflect the realities on the ground.

Humanitarian organizations rely on data to navigate highly politicized and volatile environments where resources are limited and the outcomes are not important. Data is used to make critical interventions and decisions, including whether to initiate a scale-up, allocate limited financial and human resources, and adjust ongoing operations based on their observed effectiveness. In conflict-affected regions like Somalia, high-quality data is the primary tool for learning and informing development practices that aim to be both culturally sound and technically rigorous. However, the integrity of these decisions is only as strong as the data informing them. Poor-quality data leads to poor decisions with far-reaching and often catastrophic consequences for affected communities. Inaccurate, incomplete, or cooked data can lead practitioners to underestimate the severity of a crisis or neglect priority needs.

The central problem addressed in this reflection case study is that Quality Assurance (QA) is often seen as just another box to tick in the data collection process and treated as an operational requirement rather than a strategic investment. To move beyond this paradox, the humanitarian sector must recognize that data excellence is not merely about the quantity of information, but about its validity, reliability, and integrity.

Background

Some years back, I was responsible for coordinating and leading a large data collection exercise for a humanitarian research project in Somalia. Alongside coordinating the field teams and managing client communications, I also took on responsibility for overseeing the quality assurance process, working closely with a small QA team within our project, to ensure that the data genuinely reflected the experiences and perspectives of the communities being studied.

Before field deployment, all researchers underwent comprehensive training covering research ethics, interview techniques, facilitation skills, informed consent, and data collection protocols. Detailed respondent lists were prepared, quality standards were agreed upon, initial security assessment was made and field teams received continuous technical support throughout the exercise.

Despite these preparations, the quality assurance process still uncovered several serious attempts to manipulate the data, a reminder that good training and clear protocols, on their own, are not enough to guarantee data integrity.

The Quality Assurance Process

Our quality assurance approach went beyond reviewing interview transcripts. At the time, there was no independent quality assurance body involved, we treated the review process with the same rigor an external body would have applied.

Every qualitative interview and Focus Group Discussion (FGD) was audio recorded (with participants' consent), and my QA colleagues and I reviewed each one systematically rather than through random spot checks. We cross-referenced each recording against its submitted transcript, respondent list and interview guide to check whether the questions had been asked as written, whether responses matched the respondent's recorded details, and whether the conversation felt genuine and unscripted rather than rehearsed or substituted.

This process worked, but it also revealed its own limitation, because the Quality assurance function sat within the same team delivering the project, it depended entirely on our own diligence and willingness to flag problems even when doing so reflected on our own performance and timelines. Not every organization is willing or able to commit that same level of human resources and energy to quality assurance. In my experience, many institutions invest heavily in data collection but treat Quality assurance as a lower priority. This internal review proved invaluable, but it also planted the seed for a question I would return to later: what would it take to make this kind of scrutiny truly independent, and consistently resourced, regardless of which organization is running the project?

Case 1: The Missing Focus Group Participant

One of the focus group discussions was expected to include between six and eight participants. The field team had been provided with a primary list of participants as well as a reserve list to ensure the required attendance. On the day of the discussion, only five participants attended. Rather than documenting the actual attendance and reporting the challenge, the note taker decided to become the sixth participant. Working together with the facilitator, he responded to the discussion questions while pretending to be a community participant in order to create the appearance that the minimum number of participants had been achieved.

During routine quality assurance, I listened to the audio recording. Although the note taker attempted to disguise his voice, I recognized it immediately because I had worked closely with him in different assignments. Beyond the familiar voice, several warning signs emerged. His responses were noticeably different from those of the genuine participants, they were unusually polished, consistently positive, and reflected a level of familiarity with the interview guide that ordinary participants would not be expected to have. When confronted, the field researcher initially denied any wrongdoing. However, after a detailed discussion, he admitted that he had impersonated a participant because the required number of attendees had not been reached. His intention, he explained, was to avoid reporting an incomplete focus group discussion.

Case 2: Interviewing the Wrong Person

In another instance, a researcher conducting a Key Informant Interview (KII) where he interviewed a personal friend instead of the pre-identified beneficiary. The selected respondent had already been identified and

sufficient contact information had been provided. Nevertheless, rather than interviewing the intended participant, the researcher substituted his friend while presenting the interview as if it had been conducted with the correct respondent. Again, the quality assurance review detected inconsistencies. The interviewee's voice was recognized by member of the quality assurance team who were familiar with the local context. A follow-up discussion with the researcher confirmed that the respondent was indeed not the intended participant. This case demonstrated the importance of independently verifying respondent identity rather than relying solely on submitted transcripts.

Case 3: Coaching Respondents

Perhaps the most subtle form of data manipulation involved a researcher who prepared detailed answers in advance and instructed respondents to repeat those responses during interviews. While reviewing the audio recordings, I noticed that several participants responded in an unusually structured and almost identical manner. Instead of speaking naturally from their own experiences, they appeared to be reading or recalling prepared statements. When questioned, the researcher admitted that he had coached participants before the interviews. His justification was that respondents often gave very brief answers and he wanted to ensure “better quality” responses. In reality, the practice fundamentally undermined the purpose of the research. Rather than capturing authentic experiences and opinions, the interviews reflected the researcher's own assumptions.

Case 4: GPS Was Real, the Data Was Not

During a large-scale quantitative survey covering both urban and rural communities, field teams were deployed according to the approved sampling plan with adequate logistical and financial support. Each team was assigned to specific villages to conduct household interviews using structured digital questionnaires.

As part of our routine quality assurance process, our team continuously reviewed incoming data and triangulated observations. During this process, we identified unusual patterns from one rural survey location. Further verification revealed that although the field team had travelled to the assigned village, they had remained there only few minutes to capture GPS coordinates that would create the impression that interviews had been conducted at the correct location.

To avoid duplicate GPS coordinates, individual researchers dispersed to different parts of the village, recorded GPS for their individual target interviews and then immediately returned to town. Instead of interviewing respondents, each researcher later completed the questionnaires from home using assumptions and fabricated responses before submitting the data that afternoon as though they had just returned from fieldwork.

When confronted, the field supervisor explained that the village was considered insecure and that they prioritized their safety by collecting only GPS coordinates before completing the interviews elsewhere. While protecting staff safety is always paramount, this decision fundamentally undermined the integrity of the research because the submitted data no longer represented the experiences or realities of the intended respondents.

Case 5: SurveyCTO Metadata Revealed Fabricated Interviews

On another quantitative research assignment, data were collected using SurveyCTO, which includes metadata features such as timestamps, GPS coordinates, and randomly generated background audio recordings. During enumerator training, all field researchers were informed that random background recordings could occur during interviews and were instructed to obtain respondents' consent before beginning each interview.

As part of our quality assurance process, we routinely reviewed these background recordings alongside submitted questionnaires. During one review, we encountered an interview that immediately raised concern. Instead of hearing a conversation between an enumerator and a respondent, the recording contained only the sound of keyboard typing accompanied by music playing inside a vehicle. There was no evidence that an interview had taken place.

The metadata strongly suggested that the questionnaire had been completed by the enumerator alone while sitting in a parked or moving vehicle, rather than through an actual respondent interview. When presented with the evidence, the field researcher admitted that the interview had been fabricated.

Lessons Learned

These incidents reinforced an important lesson, technology alone does not guarantee data quality but they are effective only when supported by systematic and independent verification conducted by local team who understand both the language and the context. Even experienced researchers may take shortcuts, substitute respondents, coach participants, or fabricate information, collect GPS and conduct interviews at home when faced with logistical challenges, time pressures, security issue or perceived performance expectations.

Without a serious quality assurance system, these issues could easily have gone undetected. The submitted transcripts appeared credible. Only systematic review of audio recordings, respondent verification, and detailed quality checks exposed the problems.

The experience also highlighted another reality. Several researchers later commented that this was the first assignment in which they had ever been required to submit audio recordings for qualitative interviews. In previous assignments with other intuitions, they had routinely submitted transcripts without any verification. Some openly acknowledged that the rigorous quality assurance process changed how they approached fieldwork because they knew their work would be thoroughly reviewed.

It is important to emphasize that these examples are not presented to criticize individual researchers or suggest that such practices are representative of the wider research community. Humanitarian data collection is often conducted under demanding conditions, including tight timelines, difficult access, security constraints, logistical challenges, and pressure to meet project deliverables. In such environments, even well-trained and well-intentioned field teams may deviate from agreed protocols, while a small minority may deliberately compromise data integrity. These realities underscore the need for strong quality assurance systems rather than assumptions about individual performance. Its purpose is not to police or judge researchers, but to reinforce a culture of integrity, learning, continuous improvement, and accountability across the humanitarian and development sector.

At a time when humanitarian resources are declining while needs continue to grow, every decision must be based on evidence that is accurate, credible, and independently verified.

From Experience to Innovation

These experiences ultimately inspired the creation of RAYID, an independent quality assurance and live verification firm dedicated exclusively to verifying the quality of the evidence informing decisions affecting lives.

We recognized that many organizations invest heavily in data collection but comparatively little in independently verifying whether the data can actually be trusted. Research firms are often responsible for both collecting data and assessing its quality, creating an inherent risk that quality concerns may be overlooked, corrected quietly, or underreported in order to protect organizational reputation or project timelines.

RAYID was established to provide an independent layer of quality assurance. We do not collect programme data or implement projects. Instead, we verify whether the evidence generated by others meets agreed quality standards through systematic review of field processes, metadata, audio recordings (where consented), respondent verification, spot checks, back checks and independent quality assessments.

Our objective is simple, to ensure that decisions affecting lives are informed by reliable evidence rather than assumptions, shortcuts, or fabricated data.

Conclusion

Every dataset tells a story. The question is whether that story reflects reality or merely appears convincing on paper. Independent quality assurance provides organizations with confidence that the evidence guiding their decisions is authentic, credible, and worthy of trust. For humanitarian programmes where decisions affect lives, resources, and public accountability, confidence in the data informing those decisions is not a luxury, but a necessity.

RAYID - Independent Quality Assurance, Event Verification & Evidence-based Communication Firm

Insight. Quality. Impact.

info@rayid.net · www.rayid.net · +25261000000

Mogadishu, Somalia