

Evaluation or Implementation: Crossing the Line in Rural Egypt

Alyce Abdalla
Research Coordinator
aabdalla@internetegypt.com

Nadia Zibani
Research Consultant
nzibani@internetegypt.com

Introduction

As with most application of science to the realities of everyday life, compromise is a common activity in conducting program evaluations and impact assessments. Evaluators are often not asked to participate from the beginning of the project, leaving them with no baseline information. Or they can be otherwise limited in their access to information about the project, going into the field for a short period of time with a cursory understanding of the history of the project. Alternatively, they may be forced to choose the methodology based not on what is most appropriate with regards to the question, but instead based on what is most convenient for the project implementers.

In the example of the GIRLS* pilot project in Egypt, which will be shared with you here, all of the problems just mentioned were neatly avoided.¹ As researchers, we thought we had close to the ideal in impact assessment design- we were involved in the design of the implementation from the point of conceptualization, which reflected the mutual interest of our international research organization and a highly-respected international NGO experienced in working in community development. The research experience of The Research Group* (RG) would be integrated with the implementation expertise of Child Help* (CH), and this would create a 'learning environment' in which the best possible pilot test could be conducted. The research team would have easy access to the project, and over

¹ Out of respect for our partners and concern for the pilot project, the identifying features of the implementing agency, the evaluating organization, and the pilot project have been changed. We do not expect this change to affect the substance of our analysis.

* Starred names have been changed.

time come to be familiar to the project participants and other community members, facilitating our ability to introduce different research methodologies during the evaluation process. During the evaluation we determined that qualitative work would enable us to understand behavior change at a deeper level and to be sensitive to process issues within the daily operations of the project, and we were able to adapt our research design accordingly. We expected that close collaboration with CH would also improve our quantitative data collection, as we would be able to postpone the launch of the project until completion of a thorough baseline, including two matched control sites. As implementers, CH would have regular feedback from our qualitative research, and have the evidence of a rigorous impact assessment in presenting project outcomes, pleasing donors and furthering the chances for the expansion of the project if it proved successful.

Now, almost three years later, we have learned that by avoiding problems faced by evaluators with little involvement in project design, we are forced to confront different problems resulting from our greater involvement. First, we have found that our research and implementation design as partners in the GIRLS project has distilled the structural tensions between running a project and evaluating the project's impact on the target and secondary groups. Problems range in type, some stemming from the need to accommodate the different perspectives of the organizations. Others were based on the fundamental tension between people in the field who do the implementation work and those in the field who evaluate it. Second, and unexpectedly, we faced some instances in which we felt responsible to respond to matters that are more generally understood as relating to implementation, rather than evaluation or assessment.

We hope in this analysis to discuss the dynamics of one experience of a close partnership between an evaluation and an implementation organization. Integrating evaluation and implementation in a complex pilot project is rewarding, but difficult, and we hope to strengthen the effectiveness of similar attempts in the future by analyzing our current challenges. Throughout the paper we try to give the perspectives of both evaluators and implementers, but as members of the research organization in this partnership we have a much closer understanding of the difficulties facing the evaluators than those of the implementers. Out of respect for our partners and the people involved with the project at the community level we have changed the identifying features of the organizations and the pilot project. The rest of this paper is organized as follows. We first describe the GIRLS

project and its partners, giving the context of this example. We then discuss the structural tensions involved in merging the philosophies and work activities of a research organization with that of an implementation agency. The following section deals with the difficulty in remaining strictly ‘researchers’ given our close involvement in the conceptualization and design of GIRLS. We conclude with suggestions to ease the relationship between implementation and evaluation in a partnership between two distinct organizations.

The GIRLS Project

The GIRLS project was born out of concern for rural out-of-school adolescent girls, the most disadvantaged sub-group among the 13 million adolescent boys and girls living today in Egypt. According to the 2000 Egypt Demographic Health Survey, for example, 54% of rural Upper Egyptian girls aged 13-15 are not enrolled in the formal education system, as compared to 11% of boys. A hard-to-reach group, rural girls also have restricted mobility and limited social networks, suffer from relatively poor health and malnutrition, and are at risk of early or forced marriage. Adolescence can be a time of growth and expanding opportunities, but for many girls in Upper Egypt, it is just the opposite. The objective of the GIRLS pilot project responds to this situation, and strives to empower adolescent girls and increase their life opportunities as they transition into adulthood.

Child Help and The Research Group jointly set the objective of experimenting with creating safe space in which girls could gather, learn, socialize, and play. Following a planning phase in which two additional NGOs with extensive experience in providing literacy and life-skills classes were added as implementation partners, a holistic package of interventions for out-of-school adolescent girls was launched in 2001. The Egyptian Ministry of Youth and Sports, the Ministry of Education, and the Ministry of Health and Population also agreed to contribute resources to the project. The pilot project, named GIRLS, targets out-of-school adolescent girls who are between the ages of 13 and 15 at the beginning of the intervention. Thus far, the project has reached about 200 beneficiaries in four rural villages, with two matched villages serving as control sites. The project is complex in its design, integrating curricula of three different substantive components, including sports and recreational activities for girls, and involving boys and parents in an attempt to reach those who influence girls’ lives.

Overall coordination and management of the GIRLS intervention is the responsibility of Child Help. The Research Group contributed to the program design and manages all aspects of data collection to assess the impact of this program on beneficiaries and secondary target groups such as parents, teachers of program curricula, and boys in the community. Initially, process evaluation at mid-point of the project was to be the responsibility of CH. When this activity lagged by several months and then was carried out in a somewhat cursory manner, RG felt compelled to become more involved in this aspect of data collection and analysis than had been originally planned.

Given the complex nature of the pilot project, RG designed a mix of qualitative and quantitative instruments to understand the impact of the program on its primary target group of adolescent girls. Over the course of the project, different instruments were designed as deemed necessary, and research was also conducted on secondary groups such as boys, parents, and the teachers of the project's curriculum. Baseline surveys of beneficiaries and promoters were conducted before the start of the intervention, followed by a midline survey of GIRLS beneficiaries at the half-way mark. End-line surveys of beneficiaries, girls in the control villages, and promoters are planned for 2004, two months after the end of the program, allowing for a rigorous impact assessment of the pilot project. Qualitative research methods were used to collect information from secondary target groups; with the parents of beneficiaries, for instance, focus group discussions were conducted to understand how parents view their daughters' roles in the family and community. Focus group discussions have also been conducted with adolescent boys. As a result of these multiple activities, field researchers from RG are familiar to the local community, and have time in the field for a variety of less structured research techniques such as participant observation.

Tensions Between Organizations

Both the implementation and the evaluation organizations are committed to increasing the life opportunities of out-of-school rural adolescent girls, and are working together towards that goal. In practice, however, each organization does have its distinct philosophies and professional responsibilities. In a partnership, these differences can be problematic. The following examples show how two well-regarded, professional

organizations can partner with the best of intentions, yet still face problems in working well together.

Early in our partnership, we discovered that while we might have the same goals, we take different routes to reach them. This particular example concerns RG's request to link data collected in the baseline survey to individual program beneficiaries. As researchers, we believed that giving identification numbers to the beneficiaries would respect their confidentiality while allowing for a longitudinal research design using individual level data. As implementers, CH felt that out of respect for the unique individuality of the beneficiaries, they should refuse to assign numbers to their names. Both organizations' intentions stemmed from respect for the girls in the project, but this shared respect called for different operational procedures.

As in evidence in the previous example, both Research Group staff and those at Child Help care about the beneficiaries, and hope for the success of the program. From the perspective of the researchers, however, our professional responsibility is to assess impact, using a rigorous and methodologically sound research design. The direction and magnitude of the changes we record in our data do not necessarily affect the quality of our work or how it will be evaluated by others; in fact, finding negative results could be considered evidence of our integrity to the research design in spite of a close affiliation with the program's development and operations.

As managers and implementers of the pilot project, CH staff members have a different relationship to the project. Unlike the evaluators, their professional reputations are affected by the program's success or failure, particularly for the staff in the field. Thus it is possible that 'successful' work of RG could potentially document what is perceived as failed work of CH. While they are very interested in hearing 'success stories', discussions of difficulties or failings are less appealing. Structurally, then, tension exists between RG's role and that of CH, and might be expected to affect the way in which we are able to work together. This tension between implementers and evaluators differs at various levels within the organization. At the senior level are people with a larger perspective, seeing the benefit of a tested program when deciding ways to allocate future resources. At the operation level in the field, however, the tension between researcher and implementer is more acute, where the latter may easily feel threatened by the work of the former. Accordingly, decisions made

at senior management levels of the two organizations may find easy agreement, but operations in the field can be more problematic.

Indeed, CH staff members do feel like they are under inspection when a researcher goes to the project site. This feeling is exacerbated by visits made on short notice, and by variations from the scheduled research activities. As evaluators, we think of research as a dynamic process, responding to circumstances which evolve over time. From RG's perspective, the research is targeted at evaluating the pilot project design, not the implementing agency. As such, we often underestimate their sensitivity to our research activities, and do not understand their reluctance to allow us full access to the field.

In fact, research access to the field has been increasingly contested over the course of the pilot project. Given the importance of field access in ensuring consistent, reliable, and thorough data collection, RG staff are wary of accepting any limitations, and suspicious of the reasons for restrictions. Instances in which we have investigated particular decisions of the implementation agency seem to be followed by greater obstacles to subsequent visits. Over the course of the project, CH has begun requesting more detailed itineraries and work plans, and greater access to our research materials.

In spite of our concerns about research integrity, the implementing agency does have very good reasons for full knowledge of our work plan. GIRLS' project site is in an area of Egypt with a history of security problems. As a result, state security forces keep close track of any foreigners in the governorate, out of concern for their safety. While this concern is rarely appreciated by the foreigners in question, as an international NGO with several projects in the area, CH must be sure to comply with any security requests. In Egypt more generally, gathering information is sensitive itself, and fielding survey questionnaires can be considered a security issue itself, particularly when US-based organizations such as RG and CH are involved. CH would not be able to do its work without maintaining good long-term relations with local government and security entities, and our association with GIRLS, - and therefore CH - means that our actions can affect their relationships. CH has been called in the past by security forces demanding information about an RG researcher's work in the field. If we do not inform them in advance of our activities, we might endanger their ability to respond effectively, and in a worst-case scenario, their ability to operate a wide range of programs in the governorate.

The issue of maintaining good relationships with security forces highlights the different concerns of implementers and evaluators. At the same time, there could be other reasons for which CH might limit our access to the field, resulting from the feeling that RG researchers visits are primarily inspections of their work. And from our perspective, we must be responsible to the professionalism of our work, necessitating good access to the field.

In addition to restricted access to the field, a second potential outcome of tension at the field level between implementers and evaluators is reduced effectiveness of our research findings. GIRLS is a complex project in the pilot stage, with many operational details under testing. Integrating literacy, sports, and life-skills programs into a curriculum targeted at largely illiterate girls is a difficult endeavor, as is the commitment to involve the community, particularly boys and parents, into the program. In our qualitative work with the teachers, brothers, and parents of the beneficiaries we have a trove of information that could be usefully applied to the program's activities. We do not, however, have a good way of passing along these findings to those who might make use of them. One problem is that the research is preliminary; we have not completed the analysis phase and therefore are not ready to publish results officially. CH is forced to work with different time-constraints than RG; in implementation work there is often not the luxury of taking time to reflect deeply before making decisions. It is hard for CH staff, then, to understand why we might take so long to finalize our research, particularly as the project develops and could benefit from our work. Conversely, we feel that CH is trying to rush our work. As our partner organization, CH does have a legitimate concern in questioning our reluctance to share our findings with them. In truth, we lack a practical mechanism by which we can give this information to CH in a way that protects the identities of the participants, to which we feel a basic ethical obligation. This issue is particularly difficult with qualitative findings and case study material.

On a practical level, we did not devise a mechanism early in the program by which we can share confidential information without the field staff at CH feeling criticized or attacked. Conveying some findings, for example that classroom absentee records are not always accurately kept by program teachers, would confirm their concerns that our primary objective is to inspect their daily work. As a result, our contribution to the programmatic strength of the project is diminished, as is the usefulness of our research.

As with any partnership, regular and clear communication is generally the means to avoid potential disputes. From the start of the project members of both organizations have been in contact with each other. Senior level staff members serve on a Steering Committee dealing with overall coordination of the GIRLS Project, and the two organizations also cooperate closely on a Technical Committee handling operational issues. We also communicate by phone and e-mail to deal with specific questions. Technical Advisory Group meetings are held annually, bringing in experts from around the world as well staff from all the partner organizations to share research and experience. In the first two years of project implementation, most field problems were worked out using these mechanisms. The Research Group also funded an all- project staff retreat several months into the program, in order to help build team spirit and to directly address some of the implementation problems we had observed. However, we are still distinct organizations whose employees have different professional responsibilities, academic trainings, and short-term goals; even with these communication mechanisms in place, there remains mutual misunderstandings with regards to daily activities and objectives in the field.

Tensions Within Individuals

Another result of The Research Group's close collaboration in the conceptualization and design of GIRLS is that there are instances in which individual employees feel torn between the role of researchers and their responsibilities as human beings. Involvement in the implementation of the project is clearly outside of the range of our role as defined by our professional experience and our partnership with CH. We want to raise the question here as to whether there are any circumstances that justify stepping over the boundary between these two roles. We have found two types of situations in which this problem arises. First, there have been decisions regarding the management of the program that appear to contradict the values of RG. While we are not in charge of management, our close association with CH does cause us to feel that decisions in the project do reflect on our organization. Second, there are instances where refusing to act seems unethical on a personal and humane level. In both types of circumstances the option of remaining strictly a researcher is available, but unappealing. Concrete examples of these circumstances follow below.

The GIRLS program has been underway for almost 3 years, including significant time for training of project staff. There are 24 locally recruited young women who teach the curriculum to the beneficiaries. In August 2003, three months before the end of the program, three of these teachers were fired by the program management. Given the difficulty in recruiting and retaining young women in the conservative context of rural Egypt, members of RG were surprised to hear of the event. We felt we should have been notified by CH of the decision, and also that it was somewhat of a drastic measure, given that the program will finish soon, with potentially negative consequences within the community. After contacting the implementing agency staff, we also contacted the teachers in questions. While it is certainly not RG's role to make hiring or firing decisions, this event in the program seemed to call for documentation. As partners in the design of the project, we were also interested in seeing the criteria used to fire employees. In discussion with the teachers, we were told that one teacher had been fired because she was seven months pregnant. Another had been bringing her children to work due to difficulties in finding childcare.

As with any source of information, it is difficult to know what really happened. Certainly those at the implementing agency assert that the promoters were fired because they were not working well. But in as much as the basic philosophy of GIRLS is to empower girls in their transition to adulthood, flexibility towards the variety of pressures facing working women should be part of our operational activities. As researchers, such a programmatic decision should be entirely outside of our responsibilities. Yet as people involved with the program who both respect the teachers and believe in a range of rights for working women, we felt compelled to discuss the issue further. That said, RG observes the daily activities at a distance, and we might think that being flexible is more important than other objectives of the management, such as employee moral and productivity. A meeting with the partner organizations to discuss GIRLS determined that indeed we needed guidelines giving hiring and firing criteria and processes. As a partner organization in GIRLS, we did feel an ethical obligation to raise our concerns given rumors of using pregnancy as a criterion for firing of an employee.

The second situation testing the boundaries of research roles arises from unpredictable situations in the field that call for urgent action. This example concerns a serious problem in Egypt- female genital cutting (FGC), also known as Female Genital

Mutilation and female circumcision. Estimates of the percentage of Egyptian women who undergo some form of FGC are above 90%. About half of the GIRLS beneficiaries have not been circumcised, and the part of the curriculum discusses the dangers to undergoing such a procedure.

While conducting some home visits in one of the project villages, one of the researchers from RG was approached by a project beneficiaries' mother. This researcher was a familiar person in the village, having made frequent field visits to collect qualitative data from parents. The beneficiary had been resisting her parents' decision to have her circumcised, but the mother felt sure that it was the right thing to do. The mother asked the researcher to take her daughter to the doctor outside the village to circumcise her; she offered to pay for the transportation costs of the trip.

In this situation, the researcher had a few options. She could ignore the mother, telling her to go to someone else, and remain within her role of 'researcher'. She could call the CH program manager, and relay the message. Or she could respond to the mother's request by using the trust and respect that she had built up within the community to convince the parents not to mutilate their daughter. As a researcher, remaining uninvolved was the clear choice. As a human being, remaining uninvolved was an impossible choice.

The researcher did inform the implementing agency, but also followed up by going to the family's home accompanied by a female doctor. She did succeed in convincing the girl's family not to mutilate her. In our end-line data, there may well be one extra person who has remained uncircumcised, resulting in faulty analysis that credits the GIRLS program with one extra success. Researchers are not supposed to be impact the program activities. But we doubt that anyone could suggest that the purity of the science was more important in this example than the prevention of the genital cutting.

At RG, we were never concerned that we would be tempted to bias the data upwards, showing a greater effect of GIRLS than existed in reality. We separated the qualitative and quantitative work, so that the former dealt with process evaluations and was disconnected from the baseline, midline, and end-line surveys. We have no interest in furthering the chances of the expansion of GIRLS unless it is a success. But we do have an interest in the welfare of the beneficiaries, and in certain circumstances this may affect our ability to stay within the role defined for 'researcher'.

Conclusion

At this point in the paper, it seems appropriate to say that working with GIRLS has been a rewarding experience for all of the responsible staff at RG. The concept is innovative, and gives a marginalized group a chance to learn and grow in ways often impossible in the world in which they live. Along with our partner organizations we are involved in exploring ways to expand the program beyond the four villages involved in the pilot phase. Nor do we mean to suggest that the partnership nature of the program should be structurally altered; compared to many types of interventions, the GIRLS partnership works well to create a project that has experienced implementers working with experienced evaluators.

The structural tension between implementers and evaluators is in some ways not just expected, but also an indicator of a natural relationship. Eliminating this tension while retaining the independence of each would be difficult. With that in mind, the challenge is to effectively manage the tension such that evaluators retain access to the field and implementers are able to make use of research findings.

One way to better manage this tension is to institutionalize greater clarity of the roles and responsibilities of each partner throughout the course of the pilot project. While we did this well at the beginning of the partnership, our activities adapted to changing circumstances over time, and we did not effectively redefine our responsibilities with the implementing agency. Rather than waiting for misunderstandings to cause problems, we should have decided in the partnership agreement to meet at six month intervals to clarify our responsibilities in light of any new developments in the field.

Similarly, pre-determining other mechanisms of conveying timely information in non-threatening and transparent ways could also be useful in facilitating a good working relationship. In particular, the two organizations should have decided at the beginning how and when RG would share research findings with CH. Determining these details at the start of the partnership did not seem the most urgent task at the time, but clarity at the start would alleviate some of the later misunderstandings.

More generally, the dichotomy between the training, responsibilities, and even personalities (“do-ers” vs. “thinkers”) of implementers and evaluators might not be the most effective way to create excellent projects. The long-term goal of both implementers and evaluators is the same- to help those who need it as effectively as possible; our joint success may depend on creating an environment where each better understands and respects the

work of the other. This understanding would strengthen not just the partnership, but also the quality of the work of both the evaluator and the implementer. The former would have an enhanced sense of the daily mechanisms of project activities, and the latter would be better placed to use research findings in the intervention, and potentially assist in developing good research questions in the future.

As to dealing with conflicting obligations regarding the line between evaluator and implementer, in some circumstances the people who are distanced from the daily operations are able to effectively pinpoint problems and possibly create solutions. Strict separation is not necessary to retain professional integrity; indeed it might be unethical to remain uninvolved for the good of the research but to the detriment of the beneficiaries. Social science is a series of compromises as theory is applied to reality, and being a good researcher does not supercede the responsibility to be a good human being.