



JULY 2021

MEASURING PROTECTION OUTCOMES: EMERGING EFFORTS AND NEW OPPORTUNITIES

A Results-Based Protection Briefing Paper

*The views expressed in this report do not necessarily reflect those of IRC or the Swedish International Development Agency (Sida)



INTRODUCTION

How do you stop or prevent violence in humanitarian crises? How do you know if your actions have reduced protection risk? Why is it so difficult to program for and to measure risk reduction?

In recent years, there has been a gradual shift in how the humanitarian community thinks about these questions. We know that to achieve protection outcomes, we must become better at measuring them. Although we've made progress, the humanitarian community continues to debate the best way to improve measurement of outcomes. Organizations continue to insist that more effective measurement is necessary if we are to achieve protection outcomes and if we are to strengthen how we address protection issues in humanitarian action.

This paper summarizes perspectives, experiences, and good practices from technical experts, donors, academics, policymakers, and field practitioners who worked with InterAction on Results-Based Protection (RBP) over the last year. It reflects efforts to develop a gender-based violence prevention evaluation framework (GBV PEF), and an RBP Practitioners' Roundtable that brought together practitioners from Nigeria and Iraq to discuss designing for protection outcomes. It also includes ongoing work and conversation with NGOs and donors.

We explore why measurement is an ongoing issue and focus area in RBP, what we are learning, and what we need to do to move forward.

MEASURING INTERMEDIATE RESULTS WILL HELP US MEASURE OUTCOMES

Protection outcomes—seen as reduced risk for affected communities—are the core of what we are trying to measure. We are also interested in measuring intermediate results—changes in policy, practice, attitude, knowledge, and behavior, that combined indicate progress toward a protection outcome. Measuring protection outcomes more effectively will require introducing new tools and methods and changing some fundamental ways of working. This paper examines both. Effective measurement provides evidence of the impact of results-based protection methods on communities in conflict settings. There are several reasons why measurement is crucial to the long-term effort to shift the humanitarian community's perspective on protection. Measurement will allow us to:

- Regularly examine and add to the evidence base of what reduces risk
- Ensure that the identified elements of RBP are still relevant
- Incorporate learning and innovation to promote continuous improvement
- Gather evidence to encourage wider adoption of the practices that do work.

Practitioners—both protection and non-protection—respond to concrete demonstrations of methods that could improve their work.¹ Technical specialists require an evidence base to update previous practice and guidance. Donors have expressed a desire for more examples of success that they can use to guide the way that they fund. To understand the contribution of various methods and new ways of working, we need to demonstrate that we are, in fact, contributing to reduced risk.

BARRIERS AND CHALLENGES

SHORT PROJECT CYCLES ARE NOT CONDUCTIVE TO LONG-TERM SOLUTIONS

Most of the barriers to improved measurement of protection outcomes are not new; they reflect parallel challenges in the humanitarian sector as a whole. Practitioners continue to cite time, staff, and money as the most significant barriers. Short-term project cycles, not tied to long-term programs or strategies, continue to be a barrier to measuring necessary changes in policy, behavior, or attitudes. Short funding cycles do not provide organizations with incentives to establish long-term measurement systems. The quick pace of project design often does not allow for thinking about what specifically we need to measure when it comes to protection outcomes.

Practitioners in the GBV PEF consultations experienced a “lack of freedom to maneuver” at the design stage due to quick proposal turnaround (2-4 weeks). They highlighted that the donor influenced what a project and the indicators should look like. Practitioners from the RBP Roundtable also raised concerns that existing evaluation frameworks cannot change quickly to respond to new information and analysis. For example, it can be challenging to change indicators that are developed for specific projects, often very quickly, and without the kinds of continuous analysis that can only happen throughout the life of a project.

¹ Finding from recent evaluation of InterAction’s work on RBP

These examples suggest a real need to either separate measurement processes from project design or change the parameters of project development processes to enable more outcome-oriented approaches.

WE NEED TO MEASURE BOTH SHORT AND LONG-TERM CHANGE

Many actors focus solely or primarily on long-term change. For example, they promote behavior change within a community or the development of new policy by a government or non-state actor. While many long-term changes are indeed needed to achieve protection outcomes, short- and medium-term changes can impact risk as well.

Examples of short-term changes are ensuring appropriate targeting of livelihoods interventions to reduce negative coping strategies or supporting religious leaders to negotiate to change the behavior of armed groups. Practitioners in the RBP Roundtable felt that it can be even more challenging to measure such short-term changes because it requires continuous measurement, rather than the more common practice of measurement at the end of a project.

A pathway to a protection outcome is not direct; there are frequent contextual shifts and required changes in interventions. Continuous measurement is crucial to understand what specific changes are required, as **Error! Reference source not found.** shows. We need methods that can help us to measure both short- and long-term change.

ITERATIVE PROCESS

An iterative process is an ongoing cycle of reflecting on, learning from, and adjusting actions throughout the program cycle in order to achieve results and reduce risk. This requires flexibility, adaptability, and collaboration to include the perspectives of different stakeholders.

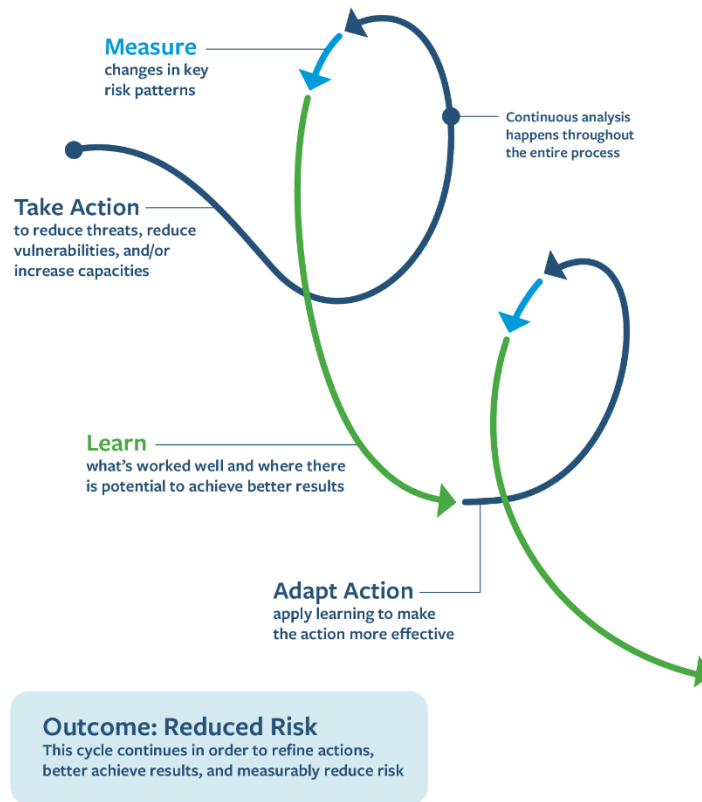


Figure 1. Measurement should be an iterative process

SOME PROGRAMS ARE INHERENTLY DIFFICULT TO EVALUATE

The question of evaluability—the extent to which results, outcomes, and the impact of a program can be measured—emerged as a critical factor in consultations during the development of the GBV PEF.² The reliance on global theories of change made it difficult to evaluate programs focusing on protection outcomes. When a program is not based on a context-specific risk analysis, it's difficult to develop a context-specific theory of the link between activities and outcomes. It is therefore difficult to evaluate those outcomes. Many programs are simply not designed to measure outcomes, and therefore, evaluation

² See [ALNAP's Evaluation of Protection in Humanitarian Action](#) for a more detailed discussion of evaluability of protection interventions

reverts to process and output measurements instead of measuring change in communities. If we are to measure protection outcomes, we need to ensure that programs are designed with outcomes as the starting point and focus. This means applying more outcome-oriented methods in the design process and using indicators that track results and outcomes, not merely project activities.

WHAT WE ARE LEARNING: OUTCOME ORIENTED METHODS

Over the past year, organizations working with InterAction have explored some existing and modified methods to determine how they can contribute to improved measurement of outcomes. Several hold promise:

- Proxy indicators and use of the risk equation
- Outcome mapping and results journals with communities
- Most Significant Change method.

Many of the methods we explored emerged from the development sector and require modification to fit into humanitarian program cycles, conflict contexts, and organizational structures. More experimenting will determine the best way to integrate these methods into existing programs.

PROXY INDICATORS AND THE USE OF THE RISK EQUATION

One of the core challenges with measuring protection is measuring incidence of risk. Whether it is due to underreporting, security risks, ethical considerations, or ineffective methods, the challenges with measuring incidence will continue to confront the humanitarian community. However, this challenge does not need to limit whether we can measure protection outcomes. One approach is to use proxy indicators to measure the changes in risk patterns identified during a context-specific protection analysis.

The *Professional Standards for Protection Work* define *risk* as “the probability of violation or threat, abuse, harm, and suffering”.¹

From a results-based protection perspective, *protection risk* is potential or actual exposure of the affected population to violence, coercion, or deliberate deprivation.

Professional Standards for Protection Work
(ICRC, 2018)

Proxy indicators are “indirect measures that are used when making direct measurements of change is not possible or appropriate.”³

The GBV PEF highlights that “proxy indicators track changes that go hand-in-hand with the change you are trying to measure. Fossil records, for example, can be used as a proxy indicator for historical climate change: we can’t directly measure what the earth’s climate was like 4,000 years ago, but the patterns of plant and animal life recorded in fossilized form can reliably tell us about it, because it goes hand-in-hand with climate change.”⁴

A protection risk has three main components, as presented in **Error! Reference source not found.** It is a result of a combined effect of the threat, the vulnerability to that threat, and the capacity to prevent, respond and recover from that specific threat.



Figure 1. Each protection risk has three components

The risk equation, so often used for protection analysis, is a fundamental tool for supporting measurement. In the field consultation portion of the GBV PEF, participants agreed that the risk equation components help think through outcome indicators that can serve as proxies for GBV risk reduction. Proxy indicators continue to hold potential for helping us better understand and capture change in a risk that itself is difficult to measure. As Figure 2 shows, the risk equation can help identify a set, or bundle, of proxy indicators that can be used to track if a risk is increasing or decreasing. This holds particular promise for

³ Corlazzoli and White (2013), pp.20-21.

⁴ [GBV PEF \(2021\)](#), p. 46

complex issues like gender-based violence, where measuring incidents is not only difficult but has particular ethical concerns.

Example risk: Adolescent boys are being recruited by a local armed group		
Threat	Vulnerability	Capacity
Local commanders require families to “give” one child per family	Adolescent boys who are not in school are the most vulnerable, as are children whose families send them out to work in the local community	Families with negotiating power and/or the means to send male household members abroad for work
Proxy indicator: Threat	Proxy indicator: Vulnerability	Proxy indicator: Capacity
Number of statements by an armed group that includes demand for boys to join the group	Rates of school attendance for adolescent boys or Percentage of families reporting that they need to send children to work	Number of households successfully preventing the recruitment of boys or Number of boys within families sent abroad to work ⁵

Figure 2. Bundling proxy indicators in line with the risk equation

Protection actors already use proxy indicators, for example, changes in attitudes towards Intimate Partner Violence (IPV) violence as a proxy for change in incidence. However, the use of proxy indicators must also be contextualized. Generalized proxies will not necessarily be effective in every context to measure change in behavior around a specific risk. The risk analysis is critical in determining context-specific proxy indicators that can be used to measure change in each context. Risk analysis should be done from the perspective of the affected population and so should the development of proxy indicators. Community members should be involved in analysis and, ideally, in developing such indicators. The more participatory the methods, the more accurate the measures of change.

⁵ While this would also be considered a negative coping strategy, it should be considered a part of risk analysis as it is a community strategy (in this fictional example).

OUTCOME MAPPING AND RESULTS JOURNALS WITH COMMUNITIES

In the past year, InterAction explored the use of outcome mapping within the GBV PEF, during the RBP Practitioners' Roundtable, and in additional workshops with NGOs. In those cases, outcome mapping was adapted to measure GBV-prevention outcomes, but there is potential for it to be used where any protection outcome has been articulated.

"Outcome mapping is a method for planning, monitoring, and evaluating projects and programs that aim to achieve lasting social and behavioral change. Outcome mapping has a range of potential uses for organizations working to prevent GBV in humanitarian contexts, including:

- It can help program teams understand complex behavior change within a community over time. This is useful for teams who want to better understand how their activities are influencing changes in the behaviors of perpetrators, vulnerable groups, and the wider community.
- It can help teams think about the pathways to change underlying their program logic. This is useful when trying to understand how the pre-conditions and underlying factors for GBV change and evolve over time.
- It is particularly useful for mapping and observing wider changes across a community, beyond the direct intended results of the program. This can help teams understand how GBV prevention activities conducted with a specific target audience can influence wider community changes beyond the direct program participants."⁶

"A results journal is a tool for collecting data about behavior change over time. What makes it a journal is the use of a community-based record of changes over time. What makes it a results journal is the focus on behavior changes within the community itself; rather than recording progress in delivering a program or set of activities. Typically, a results journal will help you to track the ultimate behavior change you are seeking to bring about, such as a reduction in IPV among migrant households in a refugee camp. But it will also help you measure the steps in the pathway to that change within the community, such as the improved awareness of IPV risk and change in underlying beliefs about IPV among the wider population."⁷

⁶ GBV PEF (2021), p.57

⁷ GBV PEF (2021), p.68

Outcome mapping typically comprises 12 steps, as illustrated in Figure 3. However, if time is limited, it can comprise only steps 1, 3, 5, and 9:

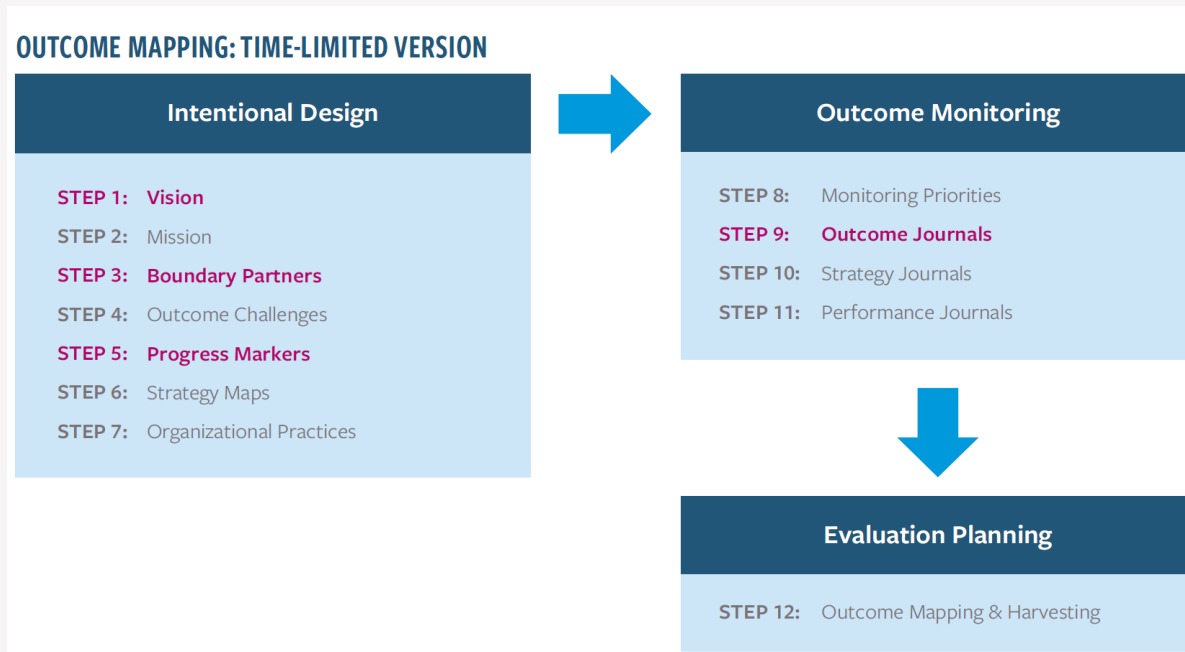


Figure 3. Outcome mapping offers a flexible approach

- **Vision** (Step 1 in Figure 3)
Describe the big-picture vision that the program (or country office) wants to achieve over the medium-term.
- **Boundary Partners** (Step 3 in Figure 3)
Choose a number of key program stakeholders, who will interact directly with the program activities (e.g., as participants in GBV-awareness-raising workshops) but who also have influence across the wider community (e.g., through involvement in women's support groups or men's social networks).
- **Progress Markers** (Step 5 in Figure 3)
Identify the key behavior changes in the community that will lead to the vision described in Step 1.
- **Results Journals** (Step 9 in Figure 3)
Design journal tools for boundary partners to use to track the changes identified in Step 5. Typically, outcome mapping tools refers to "outcome journals" rather than "results journals." We have chosen to call these Results Journals given that we want to emphasize the importance of measuring the intermediate results, such as changes in the behavior, attitudes, policy, and practice, as they relate to each component in the risk equation.⁸

⁸ [GBV PEF \(2021\)](#), p.59

Outcome mapping can use either open or closed methods of data collection; you either work with your community partners at the beginning to identify particular changes to track, or you can leave it open for them to identify the most important changes they see. Both methods have advantages and disadvantages. Closed methods can be more easily turned into quantitative tracking but create large amounts of data. Open-ended methods produce qualitative data that require a different kind of tracking and analysis and require less investment in information management and potentially time.

More broadly, open-ended data collection methods allow you to measure actual changes in the community rather than the changes that you may expect. This is even more important for intermediate results and proxy indicators, which are context-specific and, therefore, can be unpredictable. As Neil Dillon, Director of Data Conscious who helped to lead the work of the GBV PEF, shared, “If you are asking open-ended questions every week, you can track changes you may not have foreseen.”⁹ For example, you can track the level of children in school as a proxy indicator of capacity to protect them from being recruited by an armed group. But if the group changes its tactics and begins to also recruit children in school, an open-ended form of inquiry is more likely to quickly identify this change.

Outcome mapping supports you to measure both immediate and longer-term change. By measuring change continuously, you can see which short-term changes last and which do not. Another way to look at this is to make sure you measure change at different levels of an organization, as the pace of change can vary. For example, changes in behavior in armed actors on the ground might happen at a different pace than changes in policy or rules of engagement set out by their leadership.

Methods like outcome mapping are a good opportunity to demonstrate results for local actors who must demonstrate program impact but who often face shorter grant periods and higher competition for funds than INGOs. Outcome mapping can help local organizations demonstrate progress toward outcomes in the short-term. In addition, as INGOs work to adopt more of these methods, they can also work to use them with national and local partners.

⁹ Neil Dillon and his team at Data Conscious provided much of the technical support in developing the GBV PEF, as well as supporting the RBP Practitioners Roundtable.

MOST SIGNIFICANT CHANGE TECHNIQUE

"Most significant change (MSC) is a method for asking about change in the community. It is an indicator-free approach, which is typically used when you don't know what the results of your program will be, i.e., when unintended consequences are common and important. One of the core principles of MSC is to give communities an opportunity to describe how the program impacted their lives. As such, it can provide a powerful way to increase community feedback on program effects."¹⁰

The GBV PEF has adapted and developed a streamlined approach to tracking MSC with communities. It begins with identifying 3-5 broad domains the program is trying to change, such as soldiers' behavior toward civilians, or community capacity to mitigate sexual violence. Within each of those domains, the organization collects stories from community members and key stakeholders to show what they think the most significant change has been in relation to those domains identified for change. This kind of open-ended inquiry reveals the perspective of the community and those embedded in the context; they define the most significant change.¹¹ There are different ways to process these stories, as described in detail in the [GBV PEF](#) documents, including working with the community in addition to with organization staff to do a prioritization exercise and decide which stories of change are the most significant. Figure 5 shows the cycle of story selection and feedback to the places they originated.

¹⁰ [GBV PEF \(2021\)](#), p.78

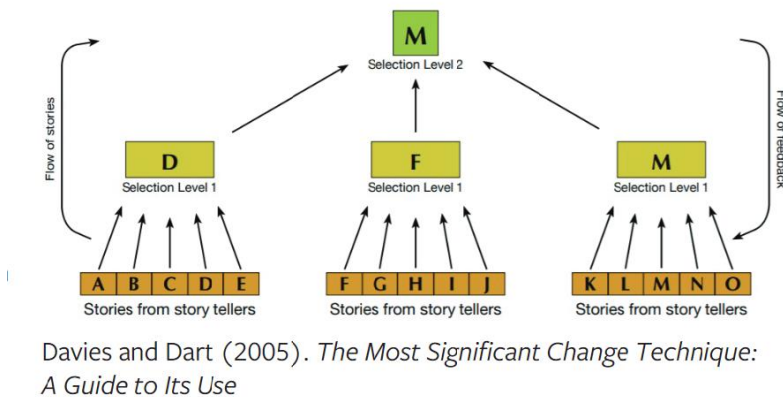


Figure 5: Most Significant Change diagram

NGOs are in early stages of testing this method. The GBV PEF field consultations in Cox's Bazar showed that some organizations were using MSC methods to measure results for reducing intimate partner violence in refugee households.

WHAT WE ARE LEARNING: NEW WAYS OF THINKING AND WORKING

WE NEED TO BE MORE OPEN TO NEW APPROACHES

For methods for measuring outcomes to be effective, we also need to create an environment that enables new and different methods of measurement to be used.¹²

One key area discussed in the RBP Practitioners' Roundtable and the GBV PEF field-level engagement was to ensure a greater focus on measuring the "world of the community" as opposed to "the world of the project." One participant pointed out that we often focus on measuring changes in attitude and behavior only in our program target group. They suggested that we need to be more open to measuring changes in the community at large, outside of the specific scope of our own programs. Measuring change is easier in a specified group or at an individual level, for example, through Knowledge, Attitude and Practice surveys, sometimes used in protection programming. We should also explore options to measure broader

¹² See [outcomes report](#) from previous Results-Based Protection Roundtable for a detailed discussion on enabling factors, including working with communities not for them, the role of creativity and flexibility, and how to simplify our tools and methods.

community change and identify the resources needed to engage in it. This also brings challenges, as some methods of measuring change at a community level may require resources that organizations do not currently have.

WE NEED TO ADOPT MORE PARTICIPATORY METHODS

Another opportunity is to look at our data collecting methods. One way to prioritize data collection is to use open-ended data collection methods, as described above using outcome mapping. Participants in the Roundtable were interested to rely on community representatives to prioritize important changes. However, they also acknowledged that it will be important to develop methods for cross-checking information; understanding the interests of various actors, including the gatekeepers of information or those who may have specific biases; and placing it in context. Robust protection analysis is always required to prioritize what changes should be measured. Linking analysis and measurement processes will help to ensure that resources are not spent gathering and analyzing extraneous information.

WE NEED TO PRIORITIZE DATA COLLECTION NEEDS

As we consider the relevant information needed to measure protection outcomes, we also need to prioritize data. Roundtable participants shared that current protection monitoring efforts often end up with more data than program teams are realistically able to analyze. It can be challenging to work with field teams to make sure there are clear understandings of what information is actually needed and what might be extraneous. One suggestion is to strengthen the collaboration and connection between MEAL teams and program teams. They can then develop clear priorities for information collection and ensure that data links to the risk equation where outcome-level change can be measured.

WE NEED TOOLS THAT WE CAN INTEGRATE WITH EXISTING MODELS

Additionally, we know that organizations are already responsible for a large amount of measurement, albeit largely tied to program deliverables and other outputs. Therefore, adding new methods to measure outcomes may strain already overloaded field teams. Moving forward, it is essential to find tools and methods that can be easily integrated into existing models of programming and measurement. The process for the design of the GBV PEF included extensive consultations with country teams to design tools that

“have the potential to fit into current practice, with minimal time or resource investment.”¹³ Experimenting with innovative methods requires investment from both NGOs and from donors. A long history of a “retreat to outputs” has also meant that individual and organizational skills are often focused on quantitative data. Organizations and donors need to work together to ensure they have capacity for qualitative in addition to quantitative methods. This means ensuring that MEAL and program staff have the necessary skills and that there are enough human resources to manage collection, flow, and analysis of qualitative data. While NGOs are interested in using resources to improve their own capacity for these methods, donors could incentivize their use by changing their requirements for monitoring and evaluation or providing additional funds explicitly for innovative methods for measuring outcomes. NGOs also emphasize the need for such funding to allow for experimentation and failure.

There is enthusiasm in the NGO community to integrate methods like outcome mapping into the way we report to donors. Donors are also enthusiastic about reporting that focuses on outcomes. The challenge is for NGOs to integrate this into monitoring efforts without simply increasing the burden on both program and MEAL teams. Once new methods are added to measure outcomes and contribution to outcomes, resources (including financial, technical, and human resources) need to be dedicated to better track the context-specific indicators produced from these methods. As we develop improved methods for measuring outcomes and demonstrate their effectiveness, there will be more opportunities for looking critically at how to optimize our available resources to best serve us, and the communities we work with.

WE NEED TO MAKE LOGICAL LINKS

There are opportunities to link analysis and measurement, as the discussion in the Roundtable demonstrated. The information collection that contributes to protection analysis often requires the same kind of trusting relationships that data collection for outcome mapping does. Practitioners shared that the activities needed for these methods of participatory measurement would fit well with existing community-based protection program activities. For example, if you regularly engage with the community or work with a community committee or group, you could integrate data collection for results journals with that regular work. More broadly, there was a general consensus that these are methods that program and MEAL staff

¹³ [GBV PEF \(2021\)](#), p.19

can do together. Some data collection methods that rely on sensitive information or pre-existing community relationships are better done by program teams. Still, MEAL teams can play a crucial role in articulating what to measure and provide support on information management and analysis. Many organizations felt this was possible but could require strengthening communication and relationships between protection program teams and MEAL staff.

WE NEED OPPORTUNITIES TO EXPERIMENT WITH DIFFERENT APPROACHES

Our movement in the direction of outcomes will rely on trial and error. To be bold, NGOs will need the security to experiment without fear of the negative consequences of failure. Donors and NGO leadership both play a key role in giving space—and resources—for failure and learning. The humanitarian community must explore specific methods that can help support improved measurement of protection outcomes, and identify key enablers that are required to support those methods. In doing so we can continue to take concrete steps toward reducing the serious risks facing communities in armed conflict and have effective means to produce the evidence that risk has been reduced.