

THE CURVE

EVIDENCE FOR DECISION-MAKING:

Incorporating Evidence into a
Responsive Feedback Approach



An Introduction to Responsive Feedback

Responsive Feedback

When a program is active in the field, decisions about program adjustments are often made quickly and expeditiously as the implementer of the program draws on experience, intuition and guess work. All of this leaves the implementer uncertain about whether steps taken to change the intervention will actually “work” and improve intervention effectiveness. A traditional Monitoring and Evaluation approach where data are gathered to assess intervention effectiveness may involve evidence gathering and interpretation too late in the program timeline to inform mid-program decisions. A complementary strategy is the **responsive feedback (RF) approach**, which is based on the premise that gathering and interpreting data throughout the intervention will help iteratively improve the probability of the success of the intervention. The capacity and capability to gather evidence is foundational to the RF approach. Evidence may be collected and used to fill gaps in knowledge and reduce uncertainty in decision-making, and improve the intervention continuously.

What is evidence and why do I need it?

Data of different types are essential for informing our programs and activities. **Evidence** is data that are systematically collected to demonstrate support (or lack thereof) for an argument or set of assumptions and leads us to accept or not accept that argument.

Types of RF evidence used in decision-making may vary by the program and its needs, resources, and scope. For example:

Ashoka University (Appendix II) used *ethnographic observations* to determine how to best support Anganwadi workers in delivering components of an iron and folic acid pill compliance intervention to pregnant mothers.

MTV Shuga, (Appendix III) used *quantitative social media data* (such as video views, watch time, and number of likes and shares) to identify the messaging strategies that resonated best with their audience.

NaijaCare, a program in Nigeria that worked with patient and proprietary medicine vendors (PPMVs), used *design workshops and consultations with organizations* to highlight key ways in which to develop a previously established program to work in a new context.

The SKY campaign, a social media campaign on tobacco control aimed at teens, used *household surveys, telephone surveys, and previously collected program data* to capture exposure to campaigns and to gain insight into how to address low program exposure in Botswana.



Why is evidence valuable for decision-making?

There may be many sources that programs draw from to make decisions—from instincts and experience to information from data collected through systematic studies. On their own, our assumptions can be valuable first steps on helping us to think through our program, but they may sometimes reflect our personal thoughts and beliefs and not fully represent the situation at hand; in short, our impressions may not always be correct. The goal of RF is to harness the power of evidence to improve decision-making to enhance the effectiveness of a program. When we can match our assumptions with objectively collected data and facts, we get closer to approaching the “truth” of the situation. This helps us to reduce bias, subject our assumptions to testing, and help us make more informed decisions.

Evidence can help to structure our decision-making in several ways. First, the systematic collection of evidence can help us to minimize our knowledge gaps about our program design or implementation. Second, once we record and review the evidence we have gathered, we can use the findings to help us determine the best course(s) of action to alter the program. Third, the clear reporting of evidence and how it has led us to identify our actions can create a transparent process so others can understand the reasoning behind our decisions (Appendix I).

The D-CIDE Framework

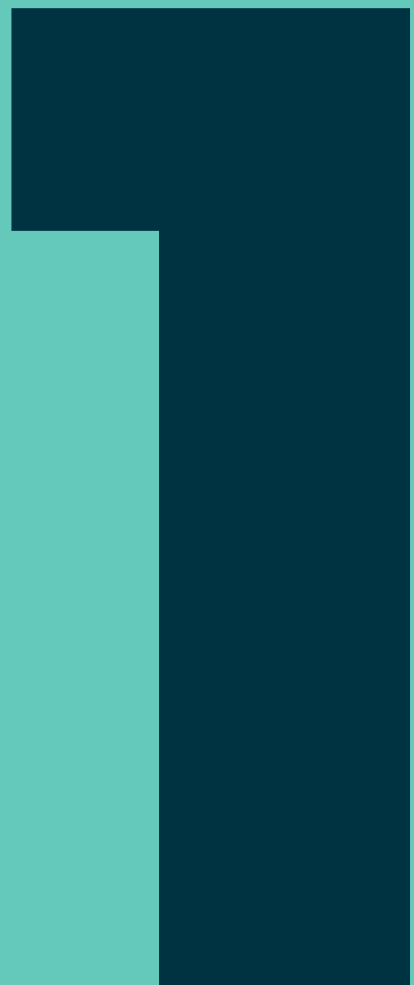
As we employ RF within our programs, some key questions include what kind of evidence is necessary, how confident we are in the data we collected, and how to use the evidence to fill potential knowledge gaps to justify actions taken to alter and adjust the intervention. In short, gathered evidence will serve the purpose of addressing knowledge gaps and presenting us with a decision: evidence may suggest a change or adjustment to the program, and we must decide whether or not to act on this evidence. The D-CIDE Framework (Figure 1) illustrates the process for using and evaluating evidence within an RF approach. Within this framework, we first define the knowledge gap we are trying to fill, then choose the correct methods to gather evidence to address this knowledge gap. As we plan to implement our chosen method(s), we take steps to ensure the quality of the evidence that we are gathering. Then, once we have gathered and analyzed RF evidence, we determine the direction or action (e.g., the change to the program) that this approach suggests, and then examine the factors (such as the programmatic environment, or context) that might impact our final decision whether or not to implement a change to our program. Through a focus on the D-CIDE framework, this document will provide the rationale for systematically incorporating evidence into RF decision-making and will provide key questions that may be used as a guide through each step (Also see Appendix I).

Figure 1- Components of the D-CIDE Framework

1. Define the knowledge gap
2. Choose correct methods
3. Inspect evidence quality
4. Determine the direction
5. Examine the context



D-DEFINE THE KNOWLEDGE GAP



D-Define the Knowledge Gap

How does RF address knowledge gaps?

A key principle of RF is to gather information that can help modify the program while it is still in the field. Modifications could be to the program design, execution, or other features that impact the effectiveness of the intervention. Often, we make decisions about programs –features and execution –when we do not have access to the full story—or when there are gaps in our knowledge about the best ways to take action to better our program. This is where evidence can be helpful in addressing such knowledge gaps—or key areas about our program where we need to know more— and reducing uncertainty about our program decisions.

There may be several ways to identify knowledge gaps in our programs. One approach may be to first identify assumptions about how our program activities lead to our program outcomes—for example, assuming that nutrition education sessions will lead to increases in nutrition knowledge. On the surface, this may seem straightforward. But as we probe this assumption, we see that there may be several other factors that we need to consider and account for so that our activities are successful—*Do participants have any challenges to attending the sessions (such as lack of transportation)? Are we advertising the sessions through the right channels so that participants are aware of these offerings? Do lessons present culturally-appropriate material that will resonate with participants?* As we think through this process, potential knowledge gaps may emerge. In other circumstances, we may already bring a problem to the table (e.g., why is attendance low at my program events) that we want to address through an RF approach.

It is useful to prioritize a knowledge gap that aligns with key program goals or areas of interest and can address areas of the program that are important for program success. Table 1 illustrates some of the key knowledge gaps addressed by RF, the potential decisions to be made, and examples.



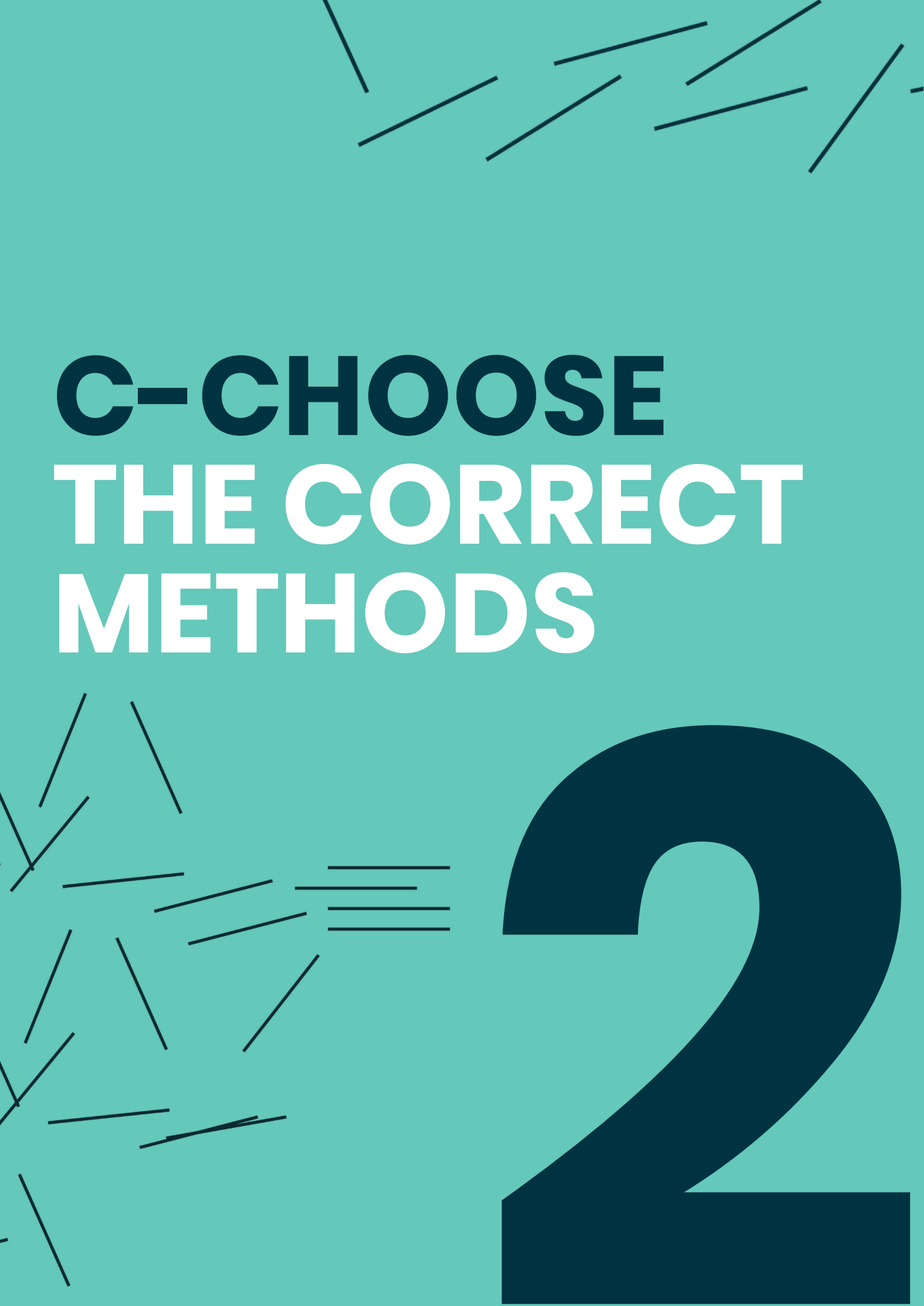
Table 1. Knowledge gaps and their related decisions

Knowledge Gap	Examples of Potential Decisions to be Made	Intervention Examples
How to develop a new approach or refine a program in its early stages.	<ul style="list-style-type: none"> • If the program is achieving the desired behavior change (or other interim outcome) or if changes are needed prior to scale up. • How to overcome potential barriers regarding initial program implementation or logistics. • If the planned intervention practices and logistics are working well in the field or if adjustments are needed. 	<p>The USAID Takamol program sought to address gender equality and female empowerment in Jordan. As this pilot program was developed, it went through several iterations, finding that content needed to be adapted to overcome barriers to engagement that became evident once the program was in the field. Tests of several approaches were undertaken until the optimal format for engagement was achieved.</p> <p>Also see: Ashoka University (Appendix II)</p>
How to apply known principles (such as those of an established program) to a new context.	<ul style="list-style-type: none"> • If an existing program will be well-received in a new setting, or if changes are needed so it resonates with the newer context. • How to customize program materials and communications to a new context and situation. • If the program must adjust to address barriers encountered from contextual factors, such as community beliefs or organizational practices. 	<p>NaijaCare was a Nigerian program conducted with PPMVs that was adapted from a program originally implemented in Kenya. Preliminary consultations with local organizations revealed that there were key concerns and regulatory considerations in Nigeria that were not accounted for in the prior version of the program. The time spent involving these stakeholders and listening to their needs to adapt, led to better buy-in over the course of the program.</p>
How to adjust if a current strategy isn't working.	<ul style="list-style-type: none"> • What underlying assumptions have been challenged by our evidence, and how to course-correct the program or underlying theory to fix this challenge. • If the feedback can stand on its own to inform program adjustments, or if more data are needed. 	<p>The Minnesota Heart Health Study used certain predetermined campaign channels to deliver their messages. However, during the program, mid-campaign surveys showed that the assumptions of the appropriate channel for the intervention were not the same across all partner communities. Researchers switched the channels in that community to increase exposure.</p>
How to gain confidence in which action to take.	<ul style="list-style-type: none"> • If there has been an area of the program that needs improvement, but there are many potential options for action (such as many potential message formats that could be used to convey a message). 	<p>The MTV Shuga Program (Appendix III) is a wide-ranging program in Nigeria that discusses family planning through a popular television show and accompanying social media content. Researchers wanted to test different messaging strategies for social media content among the target audiences for the program.</p>

Key Questions from this Section

What knowledge gap are we trying to fill?

- Prompts:
 - What areas of the program do we feel uncertain about? Are there areas in which we feel that more information is needed?
 - Are there any assumptions about how our program activities lead to our program outcomes that we should explore?
 - Are there gaps in our understanding about how to start a new project, or implement a project in a new place?
 - Are there current parts of the program that do not seem to be working, or are not working well?
 - How might addressing this knowledge gap address a key goal or outcome of our program?



C-CHOOSE
THE CORRECT
METHODS

2

C-Choose the Correct Methods

What are the different methods that can produce evidence for RF?

Methods are strategies, processes, or techniques used to gather evidence. By identifying a method or methods of data collection, we are creating the structure for systematically collecting our evidence to inform our program decisions. This systematic collection also ensures that we are conscious of the biases we hold, use well-established principles or rules in gathering the data, and are cognizant of the limitations of the data.

The method(s) that we select to guide our data collection may vary by the needs of our program. For RF, these methods center on gathering evidence while the program is still active in the field. As such, methods prioritized in this approach may differ from those that we are employing to measure the outcome of our intervention; RF-related evidence collection may need to occur in a more condensed time interval or be gathered using fewer resources. Furthermore, when selecting a method for RF, it is important to ensure that there are adequate resources to gather, analyze, and interpret evidence while the program is still in the field.

As illustrated in the table below (Table 2), there are a variety of methods that may be used to collect evidence for RF. Within the table, we present exemplar methods for gathering evidence; however there are many more methods that could be considered that are not listed here. Each has strengths that may be most useful in certain situations and fill certain types of knowledge gaps. They also have limitations that may help point to when a method may not fit the needs, budget, or scope of the program.



Table 2. Examples of methods that may be used in an RF approach

Method	Key Features	Pros	Cons
In-depth interviews	<ul style="list-style-type: none"> – Individual interviews with those who may have beneficial perspectives or information on a program. – Uses an interview guide to provide structure for questions and probes . – Responses are recorded and often transcribed. Notes may be taken. – All interviews are analyzed for themes. 	<ul style="list-style-type: none"> – Useful tool to gather data on a topic or decision point on which relatively little is known. – Provides detailed data, including the opportunity to ask follow-up questions, probe for additional information, and circle back to key questions later on in the interview. – Interviewers can monitor changes in tone and word choice to gain a deeper understanding. – Fewer participants are needed to gather useful and relevant data. 	<ul style="list-style-type: none"> – It may be time consuming to schedule interviews, conduct them, and analyze the data. – Subjective nature makes it susceptible to bias – The data may be difficult to analyze, particularly if interviews are long and many topics are covered. – Findings may not apply to a larger group of people.
Focus Group Discussion	<ul style="list-style-type: none"> – Involves the gathering of a group of people who may have beneficial perspectives on a program who are led through a guided discussion by a moderator. – The moderator uses a discussion guide to ask questions, and encourages discussion and participation by group members. – Data from focus groups are recorded and then analyzed for themes. 	<ul style="list-style-type: none"> – Provides detailed data, including the opportunity for the moderator to ask follow-up questions and probe for additional information. – Relatively inexpensive. – Individuals are more likely to provide candid responses. – Participants can build on each other’s ideas through “piggybacking”. – Can be used to understand the reasons for attitudes or behaviors in a way that cannot be captured on a survey. 	<ul style="list-style-type: none"> – Requires a skillful moderator – Requires audiotaping/ videotaping in addition to the recording of field notes. – Data may need to be transcribed for analysis. – Outspoken individuals may “hijack” and dominate a discussion. – May not be suitable for sensitive topics
Surveys	<ul style="list-style-type: none"> – Refers to the collection of information from a sample of individuals through their responses to questions. – Mostly suitable to gather quantitative data (such as gathering information through items with a preset list of answer choices), but may also be used to gather qualitative data through open-ended questions. 	<ul style="list-style-type: none"> – Relatively easy and quick to perform. – Can reach a lot of people quickly to gather information. – Could be multimodal (in-person, online, telephone) – Data can be analyzed quickly. – Respondents can have anonymity 	<ul style="list-style-type: none"> – May not provide in-depth information, particularly if quantitative items are used. – Response and return rates could be low. – Measures may need pilot-testing to check they work well for the population. – Quantitative items may vary in their quality and ability to ask the right questions (validity and reliability).
Systematic Observations	<ul style="list-style-type: none"> – Involves monitoring and noting behaviors, events or conditions as they exist in the field. – It is designed to watch or observe behavior as it organically and spontaneously unfolds in a natural environment. – Mostly suitable for gathering qualitative data. 	<ul style="list-style-type: none"> – Useful to gather data on infrequent behaviors – Useful to observe socially undesirable behaviors (topics in which people may feel less comfortable reporting their behaviors directly) 	<ul style="list-style-type: none"> – Observer’s presence may influence observed behaviors or conditions. – Requires skilled observer. – Observations may be colored by the observer’s own thoughts or experiences.

Method	Key Features	Pros	Cons
Experiments or Randomized Control Trials (RCTs)	<ul style="list-style-type: none"> – Involves assigning participants to different levels or types of treatment (conditions), and measuring the results of treatments on outcomes. – Experiments can be conducted in laboratory settings or in real-world field settings. – An experiment where participants are assigned randomly and where there is a control or comparison group is called a Randomized Control Trial (RCT). 	<ul style="list-style-type: none"> – The unique strength of experiment is its ability to link cause and effect through isolating the impact of the treatment. 	<ul style="list-style-type: none"> – Experimental research is one of the most difficult of research designs since it may be difficult to control outside factors. – Often expensive and time-consuming to conduct. – There may be ethical issues when humans are involved as subjects.
Test and Learn Cycles	<ul style="list-style-type: none"> – Involves brief tests that compare the impact of different strategies within a program. – Programs may build iteratively on the findings of these tests, using them to adjust program content as needed in the field. 	<ul style="list-style-type: none"> – Iterative approach allows for quick pivots as results are interpreted. – Can reduce uncertainty by simultaneously testing different approaches. 	<ul style="list-style-type: none"> – Results do not represent cause and effect. – This approach may require more resources, as multiple options will be tested. – Requires the ability to plan, implement, and analyze data from the test and learn cycles while the program is still in the field.
Secondary Data Analysis	<ul style="list-style-type: none"> – Uses previously collected data that can be analyzed in new ways. – May include records (such as those in a hospital or community organization), previously collected program data, or databases that have compiled information on topics of interest. – May also include the analysis of social media data that is collected on a routine basis (such as Facebook or Twitter data). 	<ul style="list-style-type: none"> – Provides a quick, low- or no-cost way of gathering information. – Benefits from the valuable work of other researchers or community members to capture relevant data. 	<ul style="list-style-type: none"> – Data are limited to what was already collected. – It may be difficult to find the right information to inform program needs.
Process Data Collection and Analysis	<ul style="list-style-type: none"> – Occurs as a part of the program to gather data on certain details of its operation and performance. – May include notes on the program's performance in the field, routine administrative data, information on the quantities or allocation of particular resources, or small interim surveys that were gathered to track intermediate indicators. 	<ul style="list-style-type: none"> – Can provide valuable data directly linked to the program in question. – Can be built into the program from the beginning, meaning it may already be linked to tracking factors that influence program outcomes. – May already be integrated into the program, and need little to no changes to provide useful information. 	<ul style="list-style-type: none"> – Process notes and other data sources may not be organized in a fashion for quick review. – Front line staff may need reassurance that analyzing interim performance data is not an audit of their performance.

The way in which each method collects evidence may make it more desirable to answer certain questions or prove/disprove certain assumptions or *hypotheses*. For example, methods such as stakeholder interviews or focus groups with participants provide a deeper understanding of the context in which we will implement the intervention. In contrast, quantitative survey data can gather information on what percentage of a target audience is performing a certain behavior.

A key feature of all methods listed is that they emphasize gathering evidence that is *collected systematically and is well-documented*. This may be particularly pertinent for RF-based evidence; results are often reviewed quickly so that decisions can be made while the program is still active in the field. Evidence that is well-documented will facilitate this quick review and will also increase our ability to illustrate to others how these findings guided our actions.

There may be situations in which more than one method can be used to gather evidence for RF decision-making. For example, combining focus groups and surveys may be a beneficial way to see how many people in our program might be performing a certain behavior (survey) and then try to understand why (focus group).

In the USAID Takamol program, several evidence sources were used in different phases of the project to gauge engagement and program acceptance with the target audience. When quantitative evidence revealed that youth (an important target group for the program) were only a small percentage of the community members who participated in program activities, staff used focus groups with youth, adaptation sessions with staff, and market research to determine how to increase engagement with this group. Throughout the process, Takamol staff also held meetings with local government and civil society organizations to reflect on what was—or was not—working, and why. Through this approach, they identified several successful strategies to increase engagement and reach among youth.

It is also important to note that RF evidence does not always need to arise from introducing a new method or data source. In many cases, it may be beneficial to assess the data that are currently being gathered in a program to determine if they can be used to inform RF activities.

In the SKY tobacco program, telephone surveys revealed that there was a low level of exposure to program content, and that these surveys were not able to reach many girls who had seen program materials. To overcome these limitations and fill knowledge gaps on why this low exposure was occurring, program staff decided to analyze the SKY Girls database, which was a database that was originally outside of the scope of the project used to gather information on teens exposed to project activities. Analysis of this database indicated that Facebook was less effective at delivering messages compared to other sources. Based on these data, program staff shifted their approach away from Facebook for future activities.

Key Questions from this Section

- What method(s) will be used to collect evidence to address this knowledge gap?
 - For each method considered:
 - What are the pros to using this method?
 - What are the cons to using this method?
- Does this method allow us to gather evidence systematically? (Can we gather the same or similar information from all data sources/participants?)
- How can we record the evidence we gather in a way that it is easy to interpret and analyze?
- Do we have the resources to analyze evidence gathered from this method while the program is still active in the field?



**I-INSPECT
EVIDENCE
QUALITY**

3

I-Inspect Evidence Quality

Why Does Evidence Quality Matter?

A key goal of the RF evidence process is to feel that we have used the evidence to minimize our knowledge gaps on implementation as much as possible. As such, we seek to *lessen* our uncertainty so that we may proceed with decisions and action. To put it another way, we may not be able to gain absolute certainty about the correctness of our decisions, but through continuing to gather evidence through our RF approach, we are seeking to gain knowledge to minimize uncertainty as much as we can. As we do so, we are seeking to structure our evidence collection and evaluation in a way that is systematic, clear, and reduces bias. Through this approach, our goal is to gather good quality evidence, regardless of the method that we choose. Gathering quality evidence can allow us to have more confidence in the results that we find from this process.

How Do I Assess the Quality of Evidence?

As we incorporate evidence into programs, gathering higher-quality evidence will allow us to make more informed program decisions. Given the broad scope of potential sources of RF evidence, there is no one marker for evidence quality. But decades of research have shown there are some core principles and tenets that serve as criteria to assess quality. Some criteria worth considering include making sure that the evidence is:

Valid- Validity ensures the accuracy of the evidence we are gathering--that we are truly measuring what we intended to measure.

Reliable- Reliability refers to if the measures produce stable and consistent results—we would want to have a measure capture information in the same way for different populations or for different participants. For research involving observations, it means that those recording the observations would report or code information in the same way (two different observers would report the same observation).

Representative. It is important to understand how our evidence does, or does not, include the range of voices that may have feedback to offer, even if it may not be timely or possible to include all viewpoints as we decide whether to act. This may require careful consideration of the identification of the sample who will be providing the evidence. Furthermore, it is useful to determine if the evidence at hand is more likely to represent one incident or the experience of one person as opposed to an experience common to many.

Credibly Sourced. When we consider an evidence source, it is valuable to ensure that we know where our data came from and that we feel this is a source of accurate information. Sources that may be of particular value are those that have prior knowledge of the subject or program, such as program staff, leaders of the members of the community, or program participants.

What is Rigor and When Does it Matter to RF?

Scientific rigor refers to using a strict, precise execution of a study design that is detailed enough to be repeated exactly. There are situations when we want the evidence to be a product of highest standards of rigor whereas in other situations, less rigorous methods will do as well. Tightly controlled conditions under which an intervention takes place helps to reduce bias, or the conscious or unconscious influencing of a study and its results. We all have bias from our own cognitions and experiences, and we may tend to generalize these thoughts as we make our program decisions. Rigor helps protect us from those biases. There are some study designs that, by the way they are structured (such as an experiment with randomized control and intervention groups), have a higher level of rigor (see Methods).

How does rigor play a role in my RF decision-making?

Rigor is an important scientific concept and provides a valuable framework for establishing cause and effect between variables. However, the most traditionally rigorous designs may not always be conducive to the conditions we are facing in the field; for example, those working to address a humanitarian crisis may not have the time and resources to conduct a study of this magnitude, or those seeking a way to structure their pilot study may want a method that provides more rich detail in their formative research with participants. As discussed in the Methods section, the nature of the program, its goals, resources, and research questions may influence the methods and study designs that may be used. Regardless of the nature of the program, we are still able to structure, gather, and document our evidence in a way that increases our confidence in the conclusions that we draw. A key in reducing bias is to create a documented trail from evidence, to decisions, to actions that can promote collective reflection.

Key Questions from this Section

Is the evidence valid?

- a. Do I feel I am measuring what I intend to measure by asking the questions that will be gathered by the selected methods?
- b. Have others within the organization (or similar organizations) used this measure before?
- c. Are there other measures that have already been tested for validity that I can use in my research?

Is the evidence reliable?

- a. Do I feel that these measures provide consistent results?
- b. Have I given clear instructions or guides for those making observations or looking for themes as they gather evidence?

Is the evidence representative?

- a. Does this evidence limit which questions were asked and which voices were heard?
- b. How broadly might this evidence represent the range of feedback I might expect on this topic?
 - a. Was it gathered in a way that considers local practices, culture and context?
 - b. Does the evidence suggest different strategies for different places or groups of people?

Is the evidence trustworthy?

- a. Where or how did I get the data to constitute evidence?
- b. Do I trust where this information came from?



D-DETERMINE THE DIRECTION



4

D-Determine the Direction

How can evidence inform directions for mid-program adjustments?

Once we have gathered evidence using our method(s) of choice, the next step is to combine, review, and analyze the evidence to determine our next steps. Analysis strategies may vary—for instance, we may look for themes (thematic analysis) within recorded interviews with participants or key informants, or look at counts or percentages from surveys. No matter what analysis strategies are used, quality RF evidence should provide a clear argument and provide information on our behavior or item of interest so we can begin to address the knowledge gap we have identified. Depending on how we have defined our knowledge gap, our gathered evidence might point to an issue—such as finding that participants are dissatisfied with a training. In other cases, it might suggest a potential solution to a problem—such as uncovering potential ways to improve program content. In either case, potential actions to address our findings may begin to emerge. Often, it is helpful to discuss these findings as a group with staff and stakeholders—such as in a pause and reflect session—to determine potential next steps for program action. This process allows for multiple viewpoints and interpretations of findings to be considered that may help us to arrive at the best course of action.



The value of mid-program evidence

Unlike outcome measures that are set at the beginning of a program and analyzed at the end, RF-based decisions require evidence sources that can provide meaningful data during the program's implementation. Typically, programs include a set of activities that we assume will lead to the achievement of the outcomes. That is, we make a number of assumptions, a theory of change, which a RF approach will allow us to test, learn and change. As we gather evidence, it may point to certain challenges present within a program or areas for improvement. For example,

The Minnesota Heart Health Program's campaign to promote health screening in three communities found out that fewer people in one of the communities were coming into the health center to get screened. Fortunately, the data gathered by the Program realized that the channel used to reach the target audience did not have sufficient reach. Specifically, two intermediate markers used by the program (brand awareness and name recall) were reported in much lower numbers **signaling that the messages were not reaching the target audience**. The program **switched channels**, from TV to direct mail, mid-course, and increased the number of people to visit the local health center for screening.

Had it been a typical M&E approach, one would not have realized the limitation of the channel until the end of the campaign. The sponsors' use of process evaluation allowed them to assess the limitations of the channel and make mid-course corrections. Let's assess the steps taken by the Sponsors here:

1. Had a clear outcome in mind: proportion of population getting screened.
2. Made some assumptions about the channels they can use to reach the target audience and encourage them to get screened.
3. Developed **intermediate goals** as markers of success of their program that pointed to program adjustments that were needed to increase message exposure.

Key Questions from this Section

- What are the key findings that this evidence has provided?
 - Does the evidence point to a potential problem we may address?
 - Does the evidence suggest a certain action to take to address a problem?
 - Does the evidence suggest we are doing something right?
 - Does the evidence provide a clear argument for action?
- How have the findings addressed our knowledge gap?
- Who should I bring to the table to interpret my findings and think through next steps?
- What action or next step do I think the findings suggest?



E-EXAMINE **THE CONTEXT**



5

E– Examine the Context

When we make decisions within programs, evidence plays one part of the final process. Evidence can serve to fill knowledge gaps and highlight a particular issue or direction that RF activities could address. However, the final decision on whether or not to implement a change to a program will rely on a number of factors, including both the magnitude of the change and features of the programmatic environment in which the change would be implemented. Given these factors, there may be different requirements for the quality or level of rigor of evidence that may be needed to justify a decision. At the end of this process, we may determine if the evidence has provided us with the confidence to proceed with a program change, or if more evidence is needed before we feel ready to make a decision.

Decision magnitude

A key feature that will set the course for information needs is the magnitude of the decision. Large-magnitude decisions may include changes to the overall program design or core program components, or costly fixes that require additional funds or the significant re-allocation of resources. Such large decisions may require a greater level of rigor (or a closer examination of quality) in our evidence. In contrast, many mid-cycle tweaks may center on smaller, easier fixes. These may include logistical considerations such as repositioning flyers or switching training rooms.

The Ashoka case (Appendix II) is an example of a potential medium- and small-magnitude set of changes. Redesigning and reprinting the information card could produce some staff time and printing expense; however, emphasizing for AWWs to refer to the card in their counseling sessions is a low-cost, relatively easy fix. The quality of evidence gathered by the ethnographers was sufficient to justify immediate action within the program.

Programmatic factors

Key factors like stakeholder preferences, cost, and urgency may lead us to our final determination on the ability of the current evidence to guide our assumptions and actions. Since RF often relies on collaboration among stakeholders, it may be important to work as a team as the decisions are reached. A process of collective reflection may be particularly valuable when the evidence is of lesser quality or if the decision is large.

The MTV Shuga case (Appendix III) illustrates how considerations were balanced to identify the part of the program to test, and how to leverage this format to gather evidence for decision-making. The television show portion of the program had a well-established format (late-stage lifecycle) that required a high cost (money, resources, and time) to test new content. However, their social media content could be quickly adapted and tested at low cost, meaning that changes to content were lower magnitude decisions that required fewer resources. Furthermore, the social media data that were gathered offered a quick way to gather and interpret how the target audience was engaging with the content.

Figure 2. Program Characteristics and Considerations

- Program goals– Key program goals may highlight priority areas of focus for RF activities and may help to guide how decisions are made.
- Lifestage of the program–Decisions made earlier in the program lifecycle, such as in prototype or pilot phases, may require less evidence to proceed.
- Urgency of the situation– Urgent crises may require the need for immediate action. It may not be possible to pursue more evidence before a decision can be made.
- Stakeholder needs–Stakeholder priorities may influence the need for evidence of a certain quality to support decisions.
- Available resources within the organization– Resources present within the organization (such as staff time, ability to analyze data, and so on) may impact the types of evidence that may be considered. This may also impact the ability to make large-scale changes to program activities.

How do we make final decisions on how to proceed?

Once all the factors are weighed with our stakeholders and partners, it may be that the evidence that is available is not sufficient to justify current action; in some cases, more evidence may be required before a decision can be made. This may be especially true for larger magnitude decisions that occur in an environment with higher stakeholder requirements or in the later stage of a program. In these cases, it may be that more evidence—particularly higher quality evidence—is needed before program action can occur. In other circumstances, the evidence provided from RF activities may guide programmatic change.

Key Questions from this Section

- How big is the decision we are making?
 - What resources will need to be acquired or shifted to implement this change to the program?
 - Does this action require the acquisition of many new resources, or the shifting of many resources (such as time, money, or staff)?
 - Does the evidence align with our current program assumptions and goals?
- What are the other programmatic considerations that may influence our decision?
 - Will stakeholders need to be involved?
 - Is the program new (such as in a pilot phase) or established?
 - Is there an urgent need for action?
 - What is the lifestage of our program?
- What is our final course of action?
 - Here are some potential conclusions from our D-CIDE process:
 - We have enough evidence to make this change to our program
 - We think this change is potentially beneficial, but we need to gather more evidence first (more information is needed to make this change)
 - We do not think this change is a good fit for our program and we must rethink our next actions (the change suggested does not fit with our program needs)

Conclusion

In summary, evidence can strengthen our ability to make confident decisions and allow us to provide some transparency in our process of making those decisions. We advocate for a decision-making process that:

- Draws from the appropriate methods that can systematically capture evidence
- Relies on quality evidence that provides clear information
- Considers the programmatic context as final decisions are reached

Appendix 1: DCIDE: A step-by-step guide for decision-making

- 1) Define the knowledge gap
- 2) Choose correct methods
- 3) Inspect evidence quality
- 4) Determine the direction
- 5) Examine context

1) Define the knowledge gap

- What knowledge gap are we trying to fill?
 - Prompts:
 - What areas of the program do we feel uncertain about? Are there areas in which we feel that more information is needed?
 - Are there any assumptions about how our program activities lead to our program outcomes that we should explore?
 - Are there gaps in our understanding about how to start a new project, or implement a project in a new place?
 - Are there current parts of the program that do not seem to be working, or are not working well?
 - How might addressing this knowledge gap address a key goal or outcome of our program?

2) Choose correct methods

- What method(s) will be used to collect evidence to address this knowledge gap?
 - For each method considered:
 - What are the pros to using this method?
 - What are the cons to using this method?
- Does this method allow us to gather evidence systematically? (Can we gather the same or similar information from all data sources/participants?)
- How can we record the evidence we gather in a way that it is easy to interpret and analyze?
- Do we have the resources to analyze evidence gathered from this method while the program is still active in the field?

3) Inspect Evidence Quality

Key Questions from this Section

Is the evidence valid?

- a. Do I feel I am measuring what I intend to measure by asking the questions that will be gathered by the selected methods?
- b. Have others within the organization (or similar organizations) used this measure before?
- c. Are there measures that have already been tested for validity that I can use in my research?

Is the evidence reliable?

- a. Do I feel that these measures provide consistent results?
- b. Have I given clear instructions or guides for those making observations or looking for themes as they gather evidence?

Is the evidence representative?

- a. Does this evidence limit which questions were asked and which voices were heard?
- b. How broadly might this evidence represent the range of feedback I might expect on this topic?
 - a. Was it gathered in a way that considers local practices, culture and context?
 - b. Does the evidence suggest different strategies for different places or groups of people?

Is the evidence trustworthy?

- a. Where or how did I get the data to constitute evidence?
- b. Do I trust where this information came from?

4) Determine the direction

- What are the key findings that this evidence has provided?
 - Does the evidence point to a potential problem we may address?
 - Does the evidence suggest a certain action to take to address a problem?
 - Does the evidence suggest we are doing something right?
 - Does the evidence provide a clear argument for action?
- How have the findings addressed our knowledge gap?
- Who should I bring to the table to interpret my findings and think through next steps?
- What action or next step do I think the findings suggest?

5) Examine the context

- How big is the decision we are making?
 - What resources will need to be acquired or shifted to implement this change to the program?
 - Does this action require the acquisition of many new resources, or the shifting of many resources (such as time, money, or staff)?
 - Does the evidence align with our current program assumptions and goals?

- What are the other programmatic considerations that may influence our decision?
 - Will stakeholders need to be involved?
 - Is the program new (such as in a pilot phase) or established?
 - Is there an urgent need for action?
 - What is the lifestage of our program?

- What is our final course of action?
 - Here are some potential conclusions from our D-CIDE process:
 - We have enough evidence to make this change to our program
 - We think this change is potentially beneficial, but we need to gather more evidence first (more information is needed to make this change)
 - We do not think this change is a good fit for our program and we must rethink our next actions (the change suggested does not fit with our program needs)

Appendix II

Ashoka University Researchers were piloting a community-based maternal education program with the goal of increasing iron and folic acid (IFA) pill compliance among pregnant women in Madhya Pradesh, India. The program relied on Anganwadi workers (AWWs) to deliver information about IFAs and their benefit through in-person counseling sessions to pregnant mothers. Researchers were interested in adding visual aides to these education sessions, but were not sure which aides might be the most effective.

Define the Knowledge Gap: Their intention was to *address knowledge gaps to inform the best ways to communicate and reinforce program content through the use of additional visual aides.*

Choose the Right Methods: The project's goal was to use evidence to determine if a counseling card, calendar, or both are successful in reinforcing the concepts in the sessions, and if adding these components to their intervention strategies was rolling out smoothly. *Ethnographic research*, involving observation of the sessions by skilled observers, was used to gauge the use of visual aides to accompany the counseling sessions, including a counseling card and/or calendar.

Inspect Evidence Quality: Ethnographers had backgrounds in health and ethnographic research (and represented a credible source for data gathering), and were briefed ahead of time by program staff about the key observations and assumptions that researchers were anticipating, with a focus on determining if the flow and content of the AWW training sessions were received well by the participants. These discussions increased the chances that ethnographers were recording information that matched the concepts the program staff wanted to capture. They observed 58 counseling sessions by 18 AWWs, so they were able to achieve a group of observations that represented many experiences using the visual aides.

Determine the Direction: There were two key observations about the counseling card used in the study: 1) that the AWWs did not show the card to the participant 25% of the time, and 2) that the counseling card did not include information about Vitamin C that was discussed in the script, indicating a disconnect between the script and the card. This evidence suggested two main fixes— to train AWWs to show the card during their sessions with mothers, and to add emphasis on the importance of Vitamin C to visual materials.

Evaluate the Context: Adding an emphasis to AWW's training to show the card was a low-cost, low magnitude decision that was able to be enacted immediately. There would be more cost involved with adapting the card, since it would take time to redesign and reprint the card. However, stakeholders reviewed the evidence and determined that it illustrated the need for adding Vitamin C to the card, so it will be included in the next wave of intervention testing.

Appendix III

MTV Shuga is a popular television and social media program in Nigeria that focuses on family planning (FP) and sexual health. The program wanted to test messaging strategies to see which strategy has the biggest impact in increasing FP communication and behaviors with their target audience. While the program has a well-established format for their television show, researchers identified their Facebook page as a promising place for testing and iteration.

Define the Knowledge Gap: Key questions emerged about what message themes and formats (such as language) may attract the most attention and would garner the most likes or shares. They had to *address the knowledge gaps of what messaging strategies will resonate best with their broad online audience.*

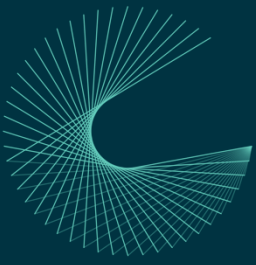
Choose the Right Methods: The goal was to create new messages and then test them to determine their performance among Facebook followers of MTV Shuga through a series of *test and learn cycles* (that is, a series of planned experiment cycles within a program that test strategies iteratively and modify program content based on the findings). Researchers wanted to use methods that would help them to first determine what message theme and language received the most engagement with viewers. The first iteration of the test focused on two potential themes identified from formative research (“sex myths” or “talk about it”) by providing a post and related video on each topic, each in four languages.

Inspect Evidence Quality: Researchers leveraged the vast amounts of social media data that was already being collected from the Facebook platform as their evidence for message engagement. For the first phase, they used the metrics of engagements, video views, and average watch times to see which message performed the best. These were measures that are already widely used and vetted by Facebook to gather quality data. The use of social media data also allowed for the potential inclusion of a wide range of audience members’ engagement behaviors within their evidence.

Determine the Direction: The English version of the “talk about it” video resonated with the widest audience—1.3 million viewers. This pointed clearly that the next phase of testing should focus on this topic and language to develop messages further in the next cycle—with English content that featured characters from the show talking to their partners about the use of protection.

Evaluate the Context: For MTV Shuga, a key decision occurred at the beginning of the test and learn cycle when they focused on generating new social media messages instead of testing content on the television show. The social media component was more suitable for low-cost fixes compared to the well-established television program which would be difficult to change—a component that was appealing to the stakeholders. Testing on social media also meant that content could be generated and changed

quickly if there was a pressing need to address a certain issue. The ability of social media to reliably gather data on key points of engagement allowed for a clear picture to emerge about the message that resonated with the audience. Through comparing these social media metrics on different posts, researchers felt this evidence gave them the confidence to identify the most successful message—and gave them a path forward to develop more content that fit within this messaging scheme.



THE CURVE