A Practitioner's Guide for

Technology Evaluation in Global Development

MIT

CITE Comprehensive Initiative on Technology Evaluation

tel technology exchange lab

About the Partners

The Comprehensive Initiative on Technology Evaluation (CITE) is the first-ever program dedicated to developing methods for product evaluation in global development. CITE is led by an interdisciplinary team at the Massachusetts Institute of Technology (MIT), and draws upon diverse expertise to evaluate products' suitability, scalability, and sustainability to develop a deep understanding of what makes products successful in emerging markets.

The Technology Exchange Lab (TEL) is dedicated to bringing communities out of poverty by driving the adoption of innovative, cost-effective and sustainable solutions that improve the lives of the least economically advantaged members of society. TEL hosts a crowdsourced database of over 650 innovations to connect practical solutions to the people who need them most. TEL also works with iNGOs and community-based organizations to identify, evaluate and implement locally appropriate solutions through co-designed development programs.

Table of Contents

An interviewee demonstrates use of water test kits in Gujarat, India. Photo by MIT CITE.

Introduction

When a person lives on less than $2 a day — as some 2.7 billion people around the world do — there is little room for a product like a solar lantern or a water filter to fail. Investing in a product that fails undermines future innovation by reducing confidence and depleting scarce resources. It is a challenge development agencies, nongovernmental organizations, and consumers themselves face every day: With so many products on the market, how do you choose the right one?

The Massachusetts Institute of Technology's Comprehensive Initiative on Technology Evaluation (MIT CITE) is the first-ever program dedicated to developing methods for product evaluation in global development. Having conducted over 12 evaluations across multiple sectors — including energy, water and sanitation, health, agriculture and more — CITE has developed a robust, versatile and replicable comparative-evaluation methodology to help organizations make better decisions around identifying and implementing products and technologies that respond to context-specific development challenges.

The Practitioner's Guide builds upon five years of research into the development of CITE's methodology in order to provide practitioners and development organizations of all sizes with practical tools and a clear framework for finding out which products and technologies are most likely to succeed in your specific development context. By following this guide, you will be able to efficiently gather the data required to drive evidence-based decisions on how best to approach your challenge area, identify potential solutions and comparatively evaluate them around relevant and context-specific criteria.

Throughout the guide, you will work through your comparative evaluation along the following roadmap as illustrated in Figure 1.

- **Identify challenges:** assess needs and define a problem statement
- **Select products:** identify potential solutions to your problem, and select products to evaluate

- **Develop evaluation criteria:** frame what matters most to you and key stakeholders as specific criteria by which to evaluate products
- **Conduct scoping study:** ramp-up your knowledge of the products and the context-specific factors of your challenge area
- **Collect data & test products:** gather data on product performance against criteria through surveying and product testing
- **Analyze data:** turn raw data into insights for decision making
- **Report results:** help your audience make evidence-based decisions around product selection and implementation

As with many aspects of running a development program, conducting product evaluations is an iterative process where you will constantly be exposed to new knowledge, perspectives and insights that will alter the course of your work. For example, you may be toward the end of evaluating five improved cookstoves when a key informant introduces you to a new product that fits squarely within your evaluation criteria. In this case, it is important to keep an open mind and recognize that this methodology is flexible and adaptable to your programmatic needs and on-the-ground realities.

The Practitioner's Guide is also modular, so that its seven steps may be followed either as a fully integrated process, or as ancillary components of your organization's existing evaluation protocols. Indeed, as you employ the CITE methodology, it is encouraged that you leverage existing institutional knowledge, resources and procedures. That way, you will be able to use the most relevant components of CITE's framework in order to strengthen your capacity to evaluate and implement development products and technologies, while also avoiding redundancy.

The Practitioner's Guide is designed for use by individual practitioners, iNGOs, development agencies, community-based organizations, and donors looking to implement innovative solutions
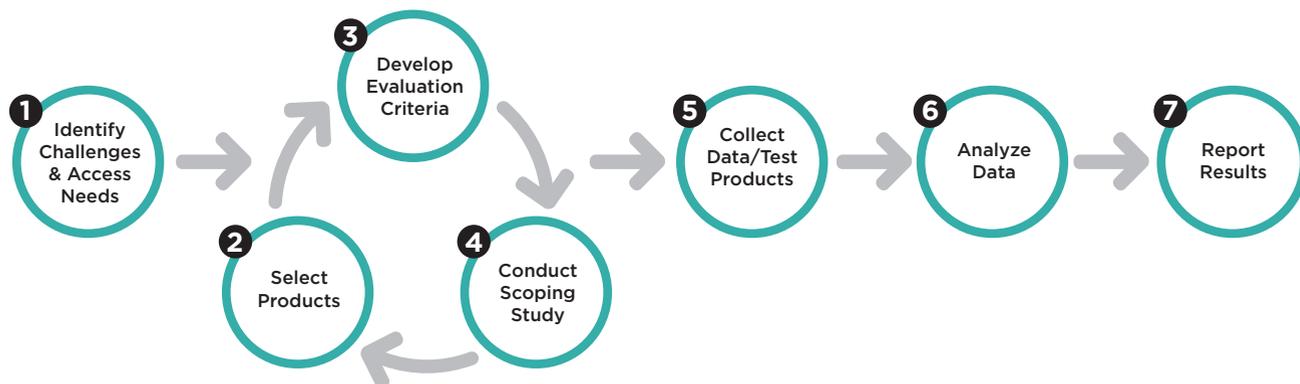
Figure 1: Comparative Evaluation Roadmap

into new or existing development programs; it is even a helpful resource for innovators, themselves, for deeper insights on how context-specific factors drive product adoption and performance. No matter what you are looking to achieve through conducting a comparative product evaluation, it is crucial to consider who this information is for, and how it will be used. Understanding the needs, priorities and motivations of your audience and key decision makers will help guide your evaluation toward answering the most important and relevant questions about the products under evaluation.

No matter the size of your organization or the resources at your disposal, we are confident that The Practitioner's Guide can offer you helpful guidance in leveraging innovative products and technologies for better development outcomes.

An interviewee demonstrates proper use of water test kits during a CITE evaluation in Gujarat, India. Photo by Jonars Spielberg, MIT CITE.

Chapter 1
Identifying challenges and assessing needs

The first step in conducting a comparative product evaluation is to have a well-articulated problem statement. In other words, what challenges are your stakeholders facing, and what opportunities are there for introducing new products or technologies to help solve these challenges?

You and your organization may already have a full understanding of the challenge at hand. For example, if an NGO with deep roots in a local community is looking to introduce point-of-use water filters in households, they may already know about user preferences and aspirations, household purchasing power, and logistical barriers to water access. Such organizations would do well to use their existing data to move toward selecting products to evaluate, while also checking for any information gaps.

For those that might be new to a specific sector or geography, it is crucial to determine the context in which you will be implementing new products and technologies by asking:

- **Who** are the key stakeholders, including beneficiaries, end-users and decision makers? How does this challenge affect them? What do they have to say about it?
- **What** are the pain points surrounding this challenge? That is, what makes this challenge so difficult to solve, and what opportunities are there for overcoming these barriers with new products and technologies?
- **Where** is this challenge occurring geographically, and how does that location affect the feasibility of potential solutions?
- **Why** is this challenge crucial to solve for your organization, and why are products and technologies an important aspect of solving this challenge?

- **When** does your timeframe begin and end?
- **How** will answering these questions provide insights for decision making, and impact decision makers?

In answering these questions, it can be helpful to conduct a baseline needs assessment by

A focus group in Bangladesh. Photo by A.M. Ahad.

interviewing key informants and experts, along with potential beneficiaries. If your organization already has a presence in target communities, you can conduct a needs assessment directly. Otherwise, it is helpful to partner with a local organization that is well connected, trusted by the community, and is able to effectively gather qualitative and quantitative data.

There are many helpful guides to conducting in-depth and rapid needs assessments, and many of them are tailored to specific sectors, such as humanitarian crisis response,[1] or community-based

energy needs.[2] Regardless of the sector you are addressing, the following activities can help lay the foundation for better understanding context-specific challenges:

Key informant interviews

Interviewing local leaders who are well-informed about their community and its members is an effective way to get a general lay of the land, while also probing deeper into more specific issues that are relevant to your program area. Key informants may be local civil or religious leaders — like school officials, village chiefs, ministers, hospital administrators, etc. — and they may be helpful in identifying further stakeholders to interview, like potential beneficiaries, households and local businesses and institutions. When interviewing key informants, consider preparing an interview guide[3] with pre-planned questions.

Target beneficiary surveys/interviews

Once you have identified target beneficiaries or households to interview, developing surveys or structured interview guides to administer to all interviewees will allow you to identify trends and gain insights around the challenges, aspirations and opportunities within the community. When conducting these interviews, be mindful of the ways in which gender, and other power dynamics affect the lives, experiences and perspectives of different interviewees.

Observation

To complement what you hear from key informants and beneficiaries, make sure you record what you see directly in the community. In this way, direct observation can let you assess basic characteristics of the community without overburdening interviewees with questions. For example, when conducting a needs assessment on the state of sanitation facilities within a school, you may prepare a checklist[4] and observe which resources are available in the community (like public latrines and water for handwashing), and what is lacking (like soap and waste disposal receptacles).

Focus groups

Focus groups are an effective, semi-structured way to get a diverse group of people to generate conversations and openly express their views on different topics. Unlike one-on-one interviews, participants may interact with each other and the moderator through a variety of activities. As with any interview, in conducting a focus group, it is important to let participants know why you are conducting the session, that there are no right or wrong answers, and that their answers and identities will remain confidential. When forming focus groups, select an appropriate moderator that will be able to effectively communicate with participants, and consider dividing participants into groups in which they will be comfortable expressing their opinions, such as by gender, occupation or position in the community.

In choosing between different methods for conducting needs assessments, it is important to harmonize your existing organizational protocols

Helpful Tip

Data Collection Apps
As a cost-effective and efficient alternative to traveling to the field for data collection, you can take advantage of communication applications designed for collecting data in low-resource settings. mSurvey[5] is an application that lets organizations conduct rapid surveys by sending out questions via SMS text messages to targeted communities and demographics in developing regions. CommCare[6] is an open source, mobile platform that enables non-technical users to build mobile apps for their frontline programs, including case management, data collection, and data management. FrontlineSMS[7] is a free open source software used by a variety of organizations to distribute and collect information via text messages (SMS). The software can work without an Internet connection and with only a cell phone and computer. Additional data-collection tools include Kobo,[8] Magpi,[9] Formhub[10] and DataWinners.[11]

with the contexts of the target regions and communities where your evaluation will take place. Above all, it is crucial to respect the will, privacy and time of those supporting your research.

Further reading — Lean Research

Humanitarian and international-development research is often conducted to understand and improve the impacts of various program interventions in the lives of populations facing poverty, vulnerability, and other challenges. Yet the impact of research activity on the lives of research subjects, communities, and local partners is often ignored. By incorporating the principles of rigor, respect, relevance and right-size into the research process, Lean Research[12] seeks to minimize burden on research subjects while maximizing the value of both the research process and outputs to stakeholders.

Case Study

Cold Storage in Mali

Since 2012, Mercy Corps has been working to empower Malian communities to cope with and recover from resource scarcity and widespread poverty. Among other challenges, the lack of energy access in Mali is a significant barrier to economic development. In order to better understand the specific challenges and hurdles related to the lack of energy access in Mali, Mercy Corps partnered with MIT D-Lab's Off-Grid Energy Group to conduct community-based energy needs assessments.

By conducting household surveys in 15 off-grid villages, and key informant interviews with local leaders and business owners, D-Lab was able to gather vital insights on the current state of energy access, the aspirational needs of community members, and existing supply chains of consumer goods within off-grid communities. Along with opportunities for expanding access to improved cookstoves and solar lighting solutions, the assessment revealed an acute challenge faced by farmers in Mali: the spoilage of vegetables due to a lack of cooling and preservation technologies.

Through subsequent research, D-Lab identified low-cost vegetable cooling and storage technologies as potentially viable solutions to this challenge. These technologies can be locally manufactured and do not require electricity, as they produce a cooling effect through the evaporation of water. CITE and D-Lab then conducted a comparative evaluation of several vegetable cooling and storage product designs available in Mali to determine which particular designs are best suited to the needs of vegetable farmers and consumers in Mali, including Mercy Corps' beneficiaries.

This research provided insights into differences in performance between individual products, unexpected benefits of the technology, and the suitability of the products for different user profiles. Additionally, based on this research, a decision making tool was created to enable potential users to determine if these products are suitable for their context, along with guidance on best practices for construction and usage of the products. Read more.[13]

Researcher installs sensors on evaporative refrigerator. Photo by Lauren McKown, MIT D-Lab.

Chapter 2
Choosing a product family and selecting products

Now that you have a better understanding of the specific challenges and social contexts at hand, you can start to explore various families of products and technologies designed to solve such challenges. For example, if you have identified an acute need among individuals with disabilities for mobility solutions, you might consider exploring wheelchairs as a product family, and begin to research multiple products within that family.

With so many development solutions and technologies emerging every day, it can seem overwhelming to choose between different product options. In order to parse down the number of

An assortment of cookstoves in Mali. Photo by Eric Verploegen, MIT D-Lab.

products you are evaluating, make a list of some of the minimum requirements — such as cost and geographic availability — that products must meet in order to achieve your goals. These requirements can not only serve as helpful boundaries to keep your evaluation right-sized, but they can also inform

Helpful Resources

Product Databases
Online resources such as the Technology Exchange Lab[14] and the Global Innovation Exchange[15] offer crowdsourced databases of products and technologies for international development. There are also product family specific databases, such as the Global Alliance for Clean Cookstoves[16] and Lighting Global.[17]

Helpful Tip

An Iterative Process
When choosing which products to evaluate, keep in mind that you will continue to discover new products throughout your scoping study. You may also be forced to rethink your assumptions regarding what product requirements are most important to end-users and other key stakeholders. In this sense, choosing products and the criteria by which to evaluate them is an iterative process that will be informed through conducting your scoping study. Also remember that for each product included in your list, additional research and testing may need to be performed. In order to contain the scope of your evaluation, consider limiting your list to 10 or fewer products.

Product Requirements

Product Options

Limits	Importance Rating			Wheelchair 1	Wheelchair 2	Wheelchair 3	Wheelchair 4
Cost $200/unit	High	**Medium**	Low	**$150**	**$80**	**$190**	**$220**
Available in country	High	Medium	**Low**	**Yes**	**Yes**	**Yes**	**Yes**
Procurement time under 3 months	High	Medium	**Low**	**2 months**	**4 months**	**6 months**	**2 weeks**
Certification by board of wheelchair professionals	**High**	Medium	Low	**Yes**	**Yes**	**No**	**Yes**
Should this product be included in evaluation?				👍 Yes	👍 Yes	👎 No	✋ Consult with decision maker

Figure 2: Selecting Products by Project Requirements

the specific criteria by which you will measure and compare products in subsequent steps of the evaluation. In determining such requirements, it is helpful to reference the opinions of existing and potential end-users, as well as sector-specific experts for insights on required product specifications. Make sure you also gather input from your team and its key decision makers to determine your organizational priorities. For example, in addition to cost and geographic availability, your organization may be bound by sourcing policies that restrict which products may be procured, and timeline for procurement.

Once you've determined these requirements, list them as in the table in Figure 2. Then determine the requirement limits (e.g., a cost limit of $200 per unit), and rank whether meeting that requirement is of high, medium or low importance. Next, you may start to fill out the table with multiple product alternatives, and list out how they meet or miss requirements.

Helpful Tip

Evaluations for Public Consumption
If you are not conducting an evaluation for a specific project or audience, but rather for general consumption — say, as a report to be disseminated to multiple development agencies — you may not have a strict list of requirements to which products must adhere. In this case, consider including a variety of products, including market leaders, emerging technologies, and products that would be widely available across different geographical markets.

Case Study

Wheelchairs in Indonesia

Over 65 million people (approximately 1 out of 117 of all people in the world) require a wheelchair for improved mobility, and in order to engage fully in productive lives. In the rush to fill this gap, however, many organizations provide products which do not meet end-user needs, and may even lead to dangerous medical complications, such as pressure sores.

In response to this challenge, CITE conducted a comparative evaluation of various wheelchair models in order to help development programs make better decisions to meet the needs of people who need a wheelchair, while also providing quality information to people with disabilities so that they can advocate on their own behalves. With literally thousands of wheelchair models on the market, and limited time and resources to conduct an evaluation, the CITE team decided to narrow the scope of their evaluation to 12 chairs. By gathering input on wheelchair requirements from its partner organization, the International Society of Wheelchair Professionals, CITE was able to select a set of 10 chairs that represents a cross-section of the most widely distributed wheelchairs in the developing world, while also representing the diversity of products available.

The team also made a conscious effort to only include wheelchair manufacturers that could produce large quantities of chairs, so the products evaluated would be accessible to global NGOs scaling programs. Read more.[18]

Wheelchair user Serafin Kangad in Mindanao, Philippines. Photo by Matt McCambridge, MIT D-Lab.

Manufacturing wheelchairs in Indonesia.
Photo by MIT D-Lab.

Once completed, determine whether to include each product within your comparative evaluation. Products that meet nearly all requirements of high and medium importance should be included in your evaluation. If a product fails to meet a strict requirement — such as wheelchair 3, which is not approved by the health ministry — it is probably not worth the time and effort to evaluate it. On the other hand, if a product meets most requirements, but misses one of medium importance or several of low importance, consider discussing this with your team and key decision makers. In the case of Wheelchair 4, decision makers might be persuaded to slightly increase the cost limit in order to accommodate such a fitting product.

Chapter 3
Framing evaluation criteria

Once you have an initial list of products to evaluate, you may start to think about what characteristics of the products are most important for you or the decision maker, and what is most important for achieving your organization's goals. Within this section, we will begin to frame the evaluation criteria around the priorities of your organization and key decision makers. After conducting your scoping study in chapter 4 — in which you will be collecting data on the priorities and preferences of additional stakeholders, such as end-users — you may refine your criteria and metrics to better reflect the context in which your evaluation will take place.

Just as you selected requirements that the products must meet in order to fit with your programmatic objectives, start to list out specific criteria, which are high level characteristics that you value in a product. Criteria may vary widely between different product families, but some general criteria might include:

- **Affordability:** how much does a product cost, and are potential purchasers and end-users willing to pay this amount? Are credit or financing mechanisms available?

- **Technical performance:** how well does a product perform its primary functions, both while being tested in a controlled setting, and within real-world settings?

- **Availability:** is this product available within your target area? If so, is it stocked in sufficient quantity to meet customer demand, and what are the lead times for replenishment? If not, what is necessary to make it accessible and available to end-users?

- **Accessibility:** Is the product carried at a location within a reasonable distance for the consumer?

- **Demand:** how complex is the value proposition of the product? Is it already in demand among potential end-users, or does it require consumer interaction to build awareness? Are manufacturers and retailers marketing the product sufficiently to create new demand?

- **Ease of use:** how easy to use is the product for end-users? Does it require training?

- **After-sales service:** how is the product maintained after it is purchased by the consumer? Are warranties or after-sale services provided? If a complex system, are spare parts and/or consumables accessible?

- **Sustainability:** how satisfied are users with the product, what advantages does the product provide, and what is the likelihood that users will continue to use it over time?

Staff at a Ugandan non-governmental organization discuss evaluation with CITE researchers. Photo by MIT CITE.

Note that each general criterion may also have more specific subcriteria, as well as metrics, which are the individual measures or observations that influence the criteria score. When selecting criteria, think about how you will measure the performance of each criterion. Whether those measurements are quantitative or qualitative, it is important to be

Single-Family Refugee Shelters

Criteria	Sub-criteria	Metrics
Cost	N/A	• Cost per-unit
Livability	Comfort	• Number of windows • Cubic feet per person at capacity
	Safety	• Door with lock y/n • Penetrable with blade y/n
Convenience	Set-up	• Time in hours to set up shelter • Number of people required for setup
	Portability	• Weight • Dimensions
Durability	Wind resistance	• Wind speed to failure
	Water resistance	• Leaks with high pressure hose y/n

Figure 3: Developing Evaluation Criteria & Metrics

able to measure each criterion to see how products compare, and to take concepts from the abstract, to the concrete.

In Figure 3, we have listed out general criteria and specific metrics by which to compare single-family refugee shelters. Assuming we have already shortlisted a set of 10 shelters that meet key requirements — e.g., capacity for a family of six, deployable to a specific region, under $10,000 per

A shelter developed using human-centered design. Photo by BetterShelter.

unit — we can further evaluate the products by selecting criteria that respond to the priorities of the decision-maker, such as end-user acceptability and logistical feasibility.

In this scenario, we want to measure how liveable shelters are, whether they are convenient to use, how durable they are, and how much they cost. A criterion such as cost is directly measured by the cost per-unit of a shelter, so there are no associated sub criteria in this example. However, subcriteria such as maintenance costs and financing mechanisms could also be taken into account. Measuring something such as liveability, however, is not as straightforward. In this case we developed two subcriteria — comfort and safety. Since these may be subjective and unquantifiable as stand-alone subcriteria, we can assign objective metrics by which to measure them, such as the number of windows and cubic meters per-person for comfort, and the presence of locks and the penetrability of the shelter with a blade for security.

When selecting criteria that are relevant to your context, make sure you have the capacity to effectively measure metrics. For example, if you wish to deploy shelters in areas with high winds and heavy rains, it may make sense to measure wind and water resistance. But measuring the wind resistance

may require testing the shelter to failure inside of a wind tunnel; a costly procedure! Testing water resistance, on the other hand, may be achieved by using a high pressure hose, or simply exposing shelters to rain.

Defining initial criteria and metrics is important because it helps limit your data collection to only answering questions that feed directly into the data analysis; it also identifies gaps in your knowledge that will need to be filled in during your evaluation,

and it informs planning for testing and protocol design. Keep in mind that the initial criterion metrics, and indeed some high-level criteria, as well, may change and need to be updated throughout the course of conducting the evaluation. Indeed, by conducting your scoping study, you will gain further insights into the criteria most valued by key stakeholders, while also learning how to assign weights to individual criteria and metrics, in order to produce a total score for each product in your evaluation.

Case Study

Evaluating Business Criteria for Scaling Malaria Rapid Diagnostic Tests

The World Health Organization estimates that nearly half of the world's population is at risk for malaria, a life-threatening, but ultimately curable and preventable disease. Malaria Rapid Diagnostic Tests (mRDTs) are an important part of the fight against the disease since over diagnosis of malaria is common across sub-Saharan Africa and non-discriminant treatment could lead to widespread resistance. However, supply chain challenges keep mRDTs from making it onto the shelves of private clinics, pharmacies and drugstores, putting patients at risk of misdiagnosis.

In 2015, MIT researchers evaluated the private sector uptake for mRDTs in Uganda where the majority of patients first seek care from private clinics, pharmacies, and drug shops. Together with the Malaria Consortium, MIT researchers conducted focus group discussions and interviews with agents in a pilot that spanned the supply chain: first-line buyers, distributors, and retailers.

By employing a Multi-Criteria Decision Analysis (MCDA) methodology, researchers identified 18 criteria that supply chain actors consider when making decisions about stocking malaria rapid diagnostic tests mRDTs. The MCDA approach

also enabled elicitation of weights to understand relative priorities and value functions to assess agents' expectations. Thus, MCDA is very useful in providing evidence regarding criteria weights and performance value to analyze data and calculate metrics, which is discussed further on in the Practitioner's Guide. Based on these findings, the MIT researchers were able to make recommendations to the private sector agents and development organizations that support them regarding business models and the design of bundled service options that increase willingness to stock mRDTs. Read more.[19]
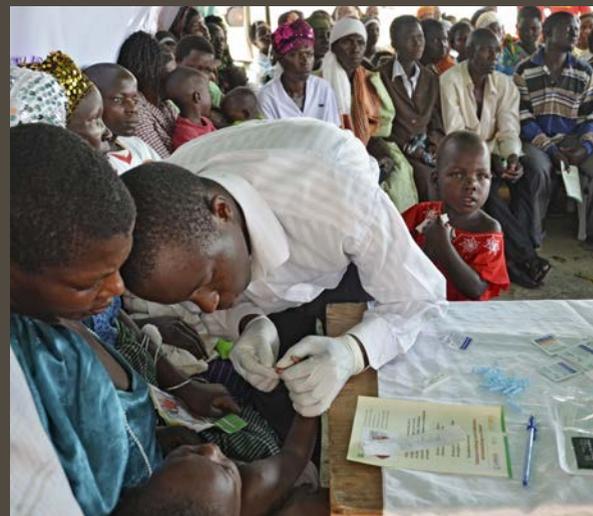
Women and children in rural Uganda look on as the service provider from a nearby health center provides malaria tests for communities on the lake's edge. © 2012 Kim Burns Case-JHUCCP, Courtesy of Photoshare.

Chapter 4
Conducting a scoping study

Before jumping into a full field evaluation, a large portion of the work can be achieved through a scoping study. Whether your scoping study is primarily desk based, field based, or a combination of the two, it is a crucial exercise in order to quickly ramp up knowledge about the product family and the context in which it is used. Scoping studies are also helpful in narrowing the scope of your evaluation, and in designing the research instruments and testing protocols necessary for answering your primary research questions.

If you haven't already done so, a field based scoping study can provide an opportunity to form partnerships with local organizations that are well connected, trusted by the community, and have an interest in the particular challenge that you have identified. They also provide a valuable opportunity to identify logistical or operational challenges that may need to be overcome in order to conduct the full evaluation. Outcomes from a scoping study help frame the full comparative evaluation but can also provide valuable insights that contribute to decision making in themselves.

Throughout this section, we will examine the following 7 components of a scoping study:

1. Expanding product knowledge

2. Mapping supply chains

3. Establishing scope

4. Defining target users

5. Mapping stakeholders

6. Understanding contexts

7. Refining criteria and metrics

While some of this information can be gleaned through a desk-based study, in chapter 5 we will introduce and discuss specific field-based data gathering tools and methods that will allow you to expand your knowledge around these 7 components, and to collect the necessary data for evaluating products.

1. Expanding product knowledge

Product evaluations are often conducted by people with expertise in the technologies being studied. If this is not the case for you or your organization, familiarizing yourself with the products and the science behind them is a helpful step in determining which product is most suitable for your project. For example, if you are looking to evaluate water filters in order to reduce arsenic levels in drinking

Workers unload food in Uganda. Photo by MIT CITE.

water, you should gain an understanding of the mechanisms by which different products filter water. While ceramic and biosand filters are effective at filtering out pathogens, they generally do not remove arsenic. In this case, technologies such as reverse osmosis, ion exchange and ultrafiltration membranes may be better suited for the project at hand.

Literature reviews are helpful in rapidly expanding your technical understand of products. Beyond simple Internet searches, a wealth of information can be found in academic articles and within institutional reports from development agencies, international NGOs, consulting firms and more. For information on specific products, manufacturer and distributor websites generally offer technical specifications and additional information on their products, along with user feedback and reviews. Throughout this process, you may be able to identify subject-matter experts on the products you are evaluation. Once you have a basic understanding of the technologies at hand, interviewing such experts can help fill gaps in your product knowledge, and also lead you to identify additional stakeholders.

2. Mapping supply chains

It is also helpful to gain a high-level view of the supply chain for your products by researching various supply chain actors, such as manufacturers, suppliers and distributors. In determining a product's cost and its potential to scale, it might be necessary to contact supply chain actors for quotes, and information on logistics and service support. Example questions to pose to a supply chain actor may include:

Helpful Resources

edX free online course in Technology Evaluation for Global Development
As a companion to the Practitioner's Guide, CITE offers an online course in Technology Evaluation for Global Development on edX, a massive open online course[20] provider. Week 4 of this course takes a deep dive into methodology, research design and applications for product scalability, including descriptions and examples of various supply-chain mapping approaches.

- What is the per-unit cost of the product?
- Is bulk ordering offered? If so, what is the price and quantity schedule?
- What are the shipping/delivery options, and the associated timeframes and landed costs for each?

- Are marketing materials and training offered?
- Are warranties, after-sale services and maintenance offered? If so, what do such warranties and services entail?
- Is there a dedicated distributor in the region?
- What quality inspections do you perform throughout the supply chain?

Sometimes, identical products from a single manufacturer are offered by various distributors, and even though such products may be equal in technical performance, the distributor service models may differ. For example, the Tulip water filter is simply a ceramic filter manufactured by Basic Water Needs, in the Netherlands. Basic Water Needs licenses the sale of their filters to distributors across developing regions, such as Soluciones Comunitarias in Central America. If you are looking to procure bulk quantities of Tulip filters for a country-wide program in Nicaragua, you might consider contacting the manufacturer directly for bulk pricing options. On the other hand, if you are looking to purchase a smaller quantity of filters for a pilot, you could avoid logistical hurdles by procuring from a country distributor, such as Soluciones Comunitarias. The latter may also come with additional benefits, such as localized after-sale service and marketing support.

3. Establishing scope

After learning more about the product family and specific products, the next step is identifying which questions you need to ask to find out if the products will meet your project's main goals. You may have already begun prioritizing evaluation focus areas while framing your criteria in section 2. Either way, it is helpful to take a big-picture view of what is most important to you, the decision-maker and key stakeholders, and to frame your main evaluation questions around those priorities. In past evaluations, CITE implemented a "3S' framework to evaluate products":

- Suitability—does a product perform its intended purpose?
- Scalability—can the supply chain effectively reach consumers?
- Sustainability—is a product used correctly, consistently & continuously over time?

By adopting a framework that responds to your organizational goals and priorities, you can pose questions specifically related to the product family, within each focus area. For example, if you are evaluation solar-powered lanterns, you might ask the following:

- Suitability—How bright is the light? How long does the battery last?
- Scalability—Where are the lanterns manufactured? How long does it take to ship them to retail outlets?
- Sustainability—Can people afford the lantern? Do people want, or have a need for a solar lantern?

Formulating these questions early in the scoping study helps to focus your evaluation, while also identifying gaps in your knowledge that will need to be addressed throughout the study.

4. Defining target users

A key part of the scoping study is to define the target end-users of the products that you are evaluating. If your project centers around a specific community, then the user cohort may be predetermined. In other instances, for example if you are evaluating products to be used for post-disaster humanitarian response, then the user cohort may differ greatly from one scenario to another. Depending on the nature of the products in question, there could be a large variation in the price

A D-Lab researcher discusses solar lanterns with a Moroccan consumer. Photo by MIT D-Lab.

and technical characteristics. Making an appropriate product choice from such a range of options is made easier by defining the target user group. For example, a $20 improved cookstove that feeds a single family might be appropriate for a community-

based development project targeting households, while a $1,000 institutional stove might be better suited for a field clinic during a humanitarian response effort.

You may have already gathered valuable data regarding end-users by conducting a needs assessment in section 1, and your organization may already have close ties with potential end-users through past and existing programs. In order to get a comprehensive profile of your target users, you can take advantage of such institutional knowledge, while also filling in any gaps by gathering the following data on target end-users:

- **Demographics:** What is their age, gender, income/standard of living and occupation or role in the household?
- **Behavior:** What do they do in a typical day? What are some productive activities and what do they do in their free time? Where do they buy products, and how much do they spend on products that fulfill a similar function?
- **Aspirations:** If they had more income, what would they spend it on? What changes would they like to see in their lives and in their communities?

By answering these questions, you can develop a deeper understanding of your target end-users, and therefore evaluate the appropriateness of different products among such users.

5. Mapping stakeholders

During your scoping study, it's important to identify all the key individuals, organizations and actors who have an interest in the success of your project, and who may have influence over the project and procurement decisions. Keep in mind that it's not always the end-users who make product decisions. Other stakeholders can have a big influence, including NGOs, governments, donors, non-

user family members, and influential community members, such as religious leaders or local advisory groups.

An easy way to make sure you include all the major stakeholders is to create a stakeholder map. Stakeholder maps are a useful tool for gaining a high level view of all the actors who may be related to your project and product, and those working on similar activities. Stakeholder mapping is especially effective as a group exercise; to begin, take a few minutes and write down all the potential stakeholders that you can think of on post-it notes or scrap paper. In Figure 4, we have listed stakeholders related to an iNGO's project for

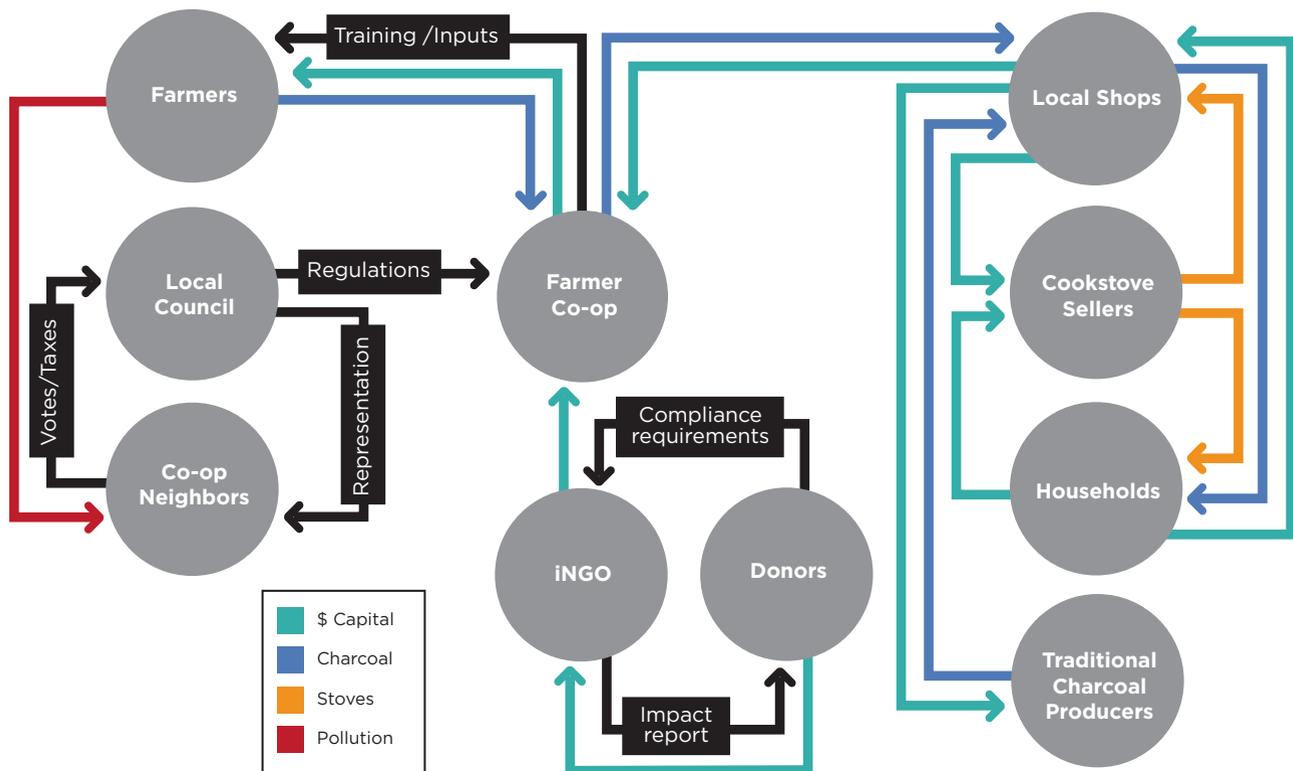### Stakeholders

Stakeholder Relationship

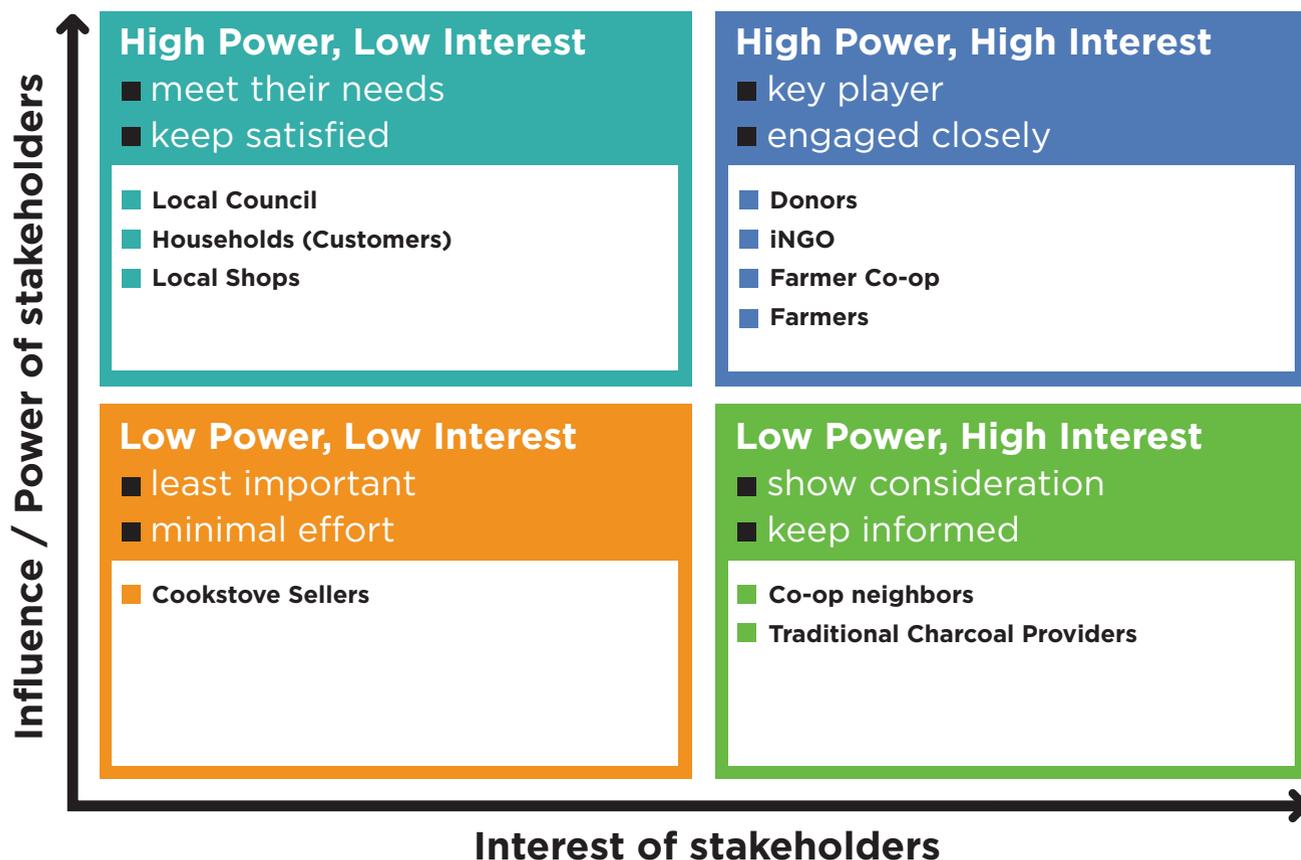**Figure 4:** Mapping Stakeholder Relationships

Figure 5: Gauging the Interest and Influence of Stakeholders

introducing green-charcoal production to a Ugandan farmer cooperative. Note that stakeholders may include those that do not benefit directly from a program — like traditional charcoal producers who are competitors — yet nonetheless form part of the local ecosystem.

Once you have listed out stakeholders, arrange them into logical clusters and then discuss and draw out the relationships, connections and transactions between stakeholders as demonstrated in Figure 4. For example, a donor provides capital to the iNGO, which provides programming and capital to the co-op. The co-op provides farmers with training on how to produce green charcoal, and through the co-op, they sell the charcoal to local businesses, who in turn sell the product to households. Through this exercise, you can not only gain a better perspective on existing relationships between stakeholders, but you can also see opportunities for creating new linkages. In this scenario, for example, it might be

possible to sell new products directly to households from producers, or through a new partnership between the co-op and cookstove sellers.

Beyond understanding the relationships between stakeholders, it is important to understand how end-users and other key stakeholders make decisions about purchasing and using new products. Do they ask for advice from family and friends? Do they want to see and touch the product before they buy it? Does having a certain type of technology change their perceived status in the community? If donors are involved, do they limit the product choices? In determining the interest and influence of multiple stakeholders, it can be helpful to create a Mendelow Stakeholder Matrix, as demonstrated in Figure 5. This can help guide your interactions with different stakeholders.

For more resources on stakeholder mapping, including an approach for Market System Mapping[24]

developed by Feed the Future Uganda and MIT's Humanitarian Response Lab, you may refer to the links in the Further Reading & Additional Resources section at the end of the Practitioner's Guide.

6. Understanding contexts

Another key part of the scoping study is to understand the context within which your products will be implemented. A context study narrows the geographic and demographic scope of the evaluation, and will vary based on the target audience of your evaluation. One important question to ask is where the results of your evaluation will be applied. Will it be within individual communities? Are these communities urban, periurban or rural? Will you implement products across multiple locations in the same country, or perhaps globally? Identifying the similarities and differences between the locations where the evaluation is being conducted, and where the results will be applied is important in gauging the relevance and effectiveness of products within different geographical contexts.

In addition to geography, your context study could include the following factors:

Environmental

What are the environmental and climatic conditions in the areas that products will be applied, and how does that affect your sector, and the performance of potential products? If you are evaluating sanitation solutions such as composting toilets, the composting performance of different toilets will vary greatly based on ambient temperature. While a given toilet model may perform well year-round in Myanmar, it may face challenges during winter months in Mongolia. Environmental differences must also be taken into account within narrower geographic boundaries. While a typical pit latrine might perform well in higher elevation areas of a given community, the same latrine could present a health hazard during the rainy season in flood-prone areas of the same community.

The local environment and lack of infrastrastructure can also affect logistics and supply chains, which are built on the foundation of physical infrastructure for transportation and storage, and require extensive communication and information infrastructure. In many developing countries, rail is not an option,

roads are in poor repair and can become unavailable during rainy seasons, and ports may have limited capacity. In addition, warehouse capacity may be limited, especially for sensitive items such as medicine requiring refrigeration. Finally, Internet and mobile phone connectivity may limit the information flows to coordinate the supply chain activities.

Social

How might social and cultural norms in your target area affect how products are perceived, adopted and used? Menstrual cups offer an affordable and long term solution for women managing their periods, and they technically function the same no matter where they are implemented. In many regions, however, the topic of menstrual hygiene management is a social taboo, and insertable products may face especially high barriers to

A researcher interviews farmers about post-harvest storage in Uganda. Photo by Jarrod Goentzel, MIT CITE.

adoption due to cultural rejection. In such cases, reusable menstrual pads may offer a more culturally accepted solution, but the effectiveness and discretion of its packaging, marketing and distribution model may still vary based on the social and cultural contexts at hand.

Political

How do local, regional and national laws and regulations factor into your product selection, and are governments and its officials supportive, adversarial or indifferent to your program? In

sectors such as health and nutrition, obtaining government approval of certain products may be necessary. In other sectors, such as energy and agriculture, there may be certain subsidies available for certain products, which can be used to your advantage. In India, for example, the national government subsidizes certain solar-powered water pump providers, presenting a clear cost advantage for selecting subsidized solar pump systems. Furthermore, In addition to the formal economy that abides by these regulations, there is an informal economy that is often more active in developing countries. This can introduce uncertainty in the market structures and restrictions in operational activities.

These are just some of the important contextual factors to consider. If your organization does not already have an established presence within your target regions and communities, consider partnering with a trusted local organization that has the capacity to provide contextual information.

7. Refining criteria and metrics

Now that you have taken a deeper dive on your evaluation through conducting your scoping study, you can leverage your expanded knowledge on products, end-users, key stakeholders and local context in order to further develop and refine your evaluation criteria and metrics. In chapter 3, you began framing your criteria around product characteristics that are important for you or the decision maker, and for achieving your organization's goals. Here — using the previous example of developing criteria for single-family refugee shelters — we can incorporate the opinions and priorities of additional stakeholders to provide a more balanced and relevant evaluation.

Single-Family Refugee Shelters

Criteria	Sub-criteria	Metrics
Cost	N/A	• Cost per-unit
Livability	Comfort	• Number of windows • Cubic feet per person at capacity
	Safety	• Door with lock y/n • Penetrable with blade y/n
Convenience	Set-up	• Time in hours to set up shelter • Number of people required for setup
	Portability	• Weight • Dimensions
Durability	Wind resistance	• Wind speed to failure
	Water resistance	• Leaks with high pressure hose y/n
Availability	Shipping time	• Time in days to ship from distributor to site
	Production capacity	• Number of shippable SKUs per month at peak production
Customization	N/A	• Number of layout options • Shelves / Hooks y/n
Compliance	N/A	• Meets local regulations y/n

Figure 6: Refining Evaluation Criteria & Metrics

Experts

What insights have you gained from product and sector specific experts, and how might that affect the importance of product specific characteristics? By interviewing a humanitarian logistician, for example, you may have learned about common logistical barriers to accessing refugee shelters. In this case, you might consider adding a criterion around accessibility or availability, and interviewing suppliers in order to measure metrics such as product delivery time and inventory output capacity.

> ### Helpful Tools
>
> **Lead User Methodology**
> Developed by MIT professor Eric von Hippel, the Lead User Methodology[25] is an approach toward product development that seeks out end-users who face needs significantly earlier than the marketplace encounters them, and who will benefit greatly from new solutions developed to address such needs. For example, manufacturers of refugee shelters hoping to develop inexpensive siding materials that can withstand extreme temperatures might look to antarctic researchers or nomadic Tuareg communities in the Saharan desert to better understand how they resolved their shelter needs, and how they may adopt and adapt such innovations at scale.

Users

Who are your target users, and how might their behaviors and aspirations impact their preferences for one product over another? By interacting with target users, you may have observed that families have multiple furniture layout arrangements and decorating styles. While this might not rank high on your organization's priorities, the ability to customize and decorate a shelter could be an important factor for maintaining the dignity and sense of wellbeing for refugee families. In such a case, you may add a criterion for customization, with specific metrics around number of layout options and the presence of features such as shelves or hanging hooks.

A Ugandan woman sews a reusable menstrual pad. Photo by Technology Exchange Lab.

Influencers

What other actors have you identified through stakeholder mapping, and how might their interests in and influence over your project affect how you or the decision-maker selects products? If you are planning to deploy for medium or long-term camps and require more permanent shelters, it may be necessary to get approval from the local authorities over issues such as zoning, shelter size and materials used. In this case, you might have a criterion such as compliance, with a straightforward metric on whether or not shelters meet local regulations. In refining your evaluation criteria and metrics, make sure you are only answering questions that are relevant to your evaluation and main research questions. The activities required to measure certain metrics may be intensive, so adding superfluous metrics to your evaluation could require an unnecessary amount of effort and resources. Through past evaluations, CITE has found that between 7–12 metrics are appropriate for a comprehensive evaluation.

Chapter 5
Gathering data

A. Preparing research instruments and protocols

In addition to desk studies, there are many methods available for gathering the data necessary in order to answer your research questions, including structured and semi-structured qualitative and quantitative surveys, direct observation and by conducting product testing experiments in the lab, field or both. The methods you choose depend greatly on the purpose of your evaluation, resources available and the priorities of your organization and its decision makers. In many cases, the required data collection methods will be readily apparent, especially when considering the specific criteria and metrics you are evaluating. For example, by referencing the metrics we developed for evaluating single-family refugee shelters in the Figure 7, it is obvious that gathering data on straightforward metrics, such as the presence of a door lock, can easily be achieved through a desk study, rather than through technical testing or interviews, which would be superfluous. For other metrics that may be more subjective, such as the number of people required for assembly, it may be helpful to gather data from a number of sources to cross check for accuracy.

Metrics for single-family refugee shelters	Desk study	Key-informant interview	End-user	Technical testing
A Cost per-unit	☑	☑		
B Number of windows	☑			
C Cubic feet/person at capacity	☑		☑	☑
D Door with lock (y/n)	☑			
E Penetrable with blade (y/n)	☑		☑	☑
F Time to set-up (hours)	☑	☑		☑
G Number of people required to set-up	☑	☑	☑	
H Weight	☑			☑
I Dimensions	☑			☑
J Wind speed to failure				☑
K Leaks with high pressure hose (y/n)				☑
L Product delivery time	☑	☑		
M Production capacity	☑	☑		
N Number of layout options	☑		☑	
O Number of shelves/hooks	☑		☑	
P Meets local regulations (y/n)		☑		

Figure 7: Determining Data Collection Method by Metric

If you are conducting an evaluation for general consumption — such as a consumer reports-style evaluation or an academic paper — it may be necessary to adhere to strict testing and research protocols. No matter the intended audience, it can be helpful to gain insights on products'

Eshita Dayani and Jonars Spielberg from MIT CITE interview a local resident about his water quality and water filter usage in Ahmedabad, India. Photo by Sydney Beasley, MIT CITE.

performance and user perceptions in real-world settings by interviewing existing product users, and by conducting field testing with existing or potential users. Of course, such activities may require additional resources in order to travel to areas with existing users and to set up field-based test beds. In determining which data collection methods are best for you, it's important to balance the rigor required for your evaluation with the resources and time at your disposal. As a starting point, you can consider the following data collection methods and tools, along with the work required to effectively prepare them and put them to use.

Surveys and structured interviews

Performing end-user surveys and structured interviews — whether before product usage, after or both — can help you understand how users perceive different products, and how they may interact with them in the future. In this sense, user surveys can be particularly helpful for gather data on criteria such as affordability, demand, ease of use, sustainability and more. Administering surveys to other stakeholders, such as suppliers and distributors, can also help

you gather supply-chain data on affordability and availability. In designing surveys and structured interview guides, it is important to take the following considerations into account:

Sample size
The number of people you survey will have an impact on the statistical significance of your findings. In order to strike a balance between available resources and rigor, make sure you get a representative impression of the community or user group at hand, without overburdening individual interviewees with unnecessary questions. In some cases, the size required for a quality sample will be readily apparent. For example, if you are trying to ascertain the availability of sanitary napkins in a peri-urban community with 20 retail shops, surveying two mom-and-pop stores may not yield a representative sample of inventories in that community. On the other hand, interviewing ten or more mom-and-pop stores along with several pharmacies will provide you with a more holistic view. Determining your sample size will also depend largely on the requirements of your organization and key decision-makers, such as funders, who may have specific guidelines for meeting statistical significance.

Survey design
In designing your surveys, it is crucial to only gather data that is directly related with your evaluation questions at hand. A good way to achieve this is by mapping all of your survey questions to your specific criteria, subcriteria and metrics, as demonstrated in Figure 8. By using the previous example of single-family refugee shelters, we could develop survey questions to gather data on our "convenience" criteria and "set-up" sub criteria by measuring metrics such as time required to set up shelter in hours, and number of people required to set up the shelter. This can be directly mapped to questions asked to a variety of interviewees, including trained personnel, untrained personnel, individuals who set up a shelter recently, and individuals who set one up further in the past. It is also important to avoid asking questions that could be answered through means other than surveying, even if they do map directly to your metrics. For example, asking end-users how much their shelter weighs in order to satisfy the "portability" subcriteria will most likely yield inaccurate results, which could be better obtained through a desk study or technical testing.

Potential Survey Questions	Answer format	Metric(s)
Did you set up the shelter yourself?	(y/n)	**F** Time to set-up **G** Number of people required to set-up
When was this shelter constructed? (date)	(date)	**F** Time to set-up **G** Number of people required to set-up
How many people did it take to set up the shelter? (number)	(number)	**G** Number of people required to set-up
How long did it take to set up the shelter? (time [hours])	(time [hours])	**F** Time to set-up
How many people are living in this shelter? (number)	(number)	**C** Cubic feet/person at capacity
When thinking about the people living in this shelter, is there enough space for them:		
During the day	(Too small, Enough space, Too much space)	**C** Cubic feet/person at capacity
At night	(Too small, Enough space, Too much space)	**C** Cubic feet/person at capacity
Does the shelter have adequate storage areas for personal belongings?	(Too small, Enough space, Too much space)	**O** Shelves/hanging hooks
Does the storage include hooks?	(y/n)	**O** Shelves/hanging hooks
Does the storage include shelves?	(y/n)	**O** Shelves/hanging hooks
What other storage options are there?	(qualitative)	**O** Shelves/hanging hooks
Is the interior layout customizable?	(y/n)	**N** Layout options
How many interior layout options are there?	(number)	**N** Layout options
Do you think the walls can be torn with a knife/blade?	(y/n)	**E** Penetrable with blade
How satisfied are you with this shelter overall?	(Not at all, Somewhat, Very Satisfied)	No direct metric
If you could, what would you change about this shelter?	(qualitative)	No direct metric

Figure 8: Mapping Survey Questions to Metrics

Format

How you physically collect your survey data is also worth considering. Paper-based surveys are less expensive and may be easier for some data collectors to manage, but transcribing paper-based surveys for digital analysis can be cumbersome, and also leaves room for error. Tablet-based surveying, on the other hand, allows for automatic data storage, analysis and visualization, yet it is more expensive and may make it difficult for taking notes during qualitative surveys. In some scenarios, tablets can also be distracting for interviewees. In deciding on survey format, it is important to factor in your organization's existing protocols and resources.

Consent & data protection

At all times, it is crucial to let interviewees know why you are requesting to interview them, to obtain their consent and to clearly communicate how data — especially pertaining to personal and/or private information — will be used. If you are collecting sensitive personal data, you must ensure that you are able to protect the security and privacy of the data whether it is in a physical or digital format.

Case Study

Single-Parameter Water Test Kits

According to the United Nations, 94% of India's population has access to an improved water supply. However, using an improved water source does not guarantee that water is safe, and while 91 million in India may be "water poor," many more are likely "safe-water poor." In areas without access to reliable and affordable testing labs, portable water test kits offer a viable solution. From 2014 to 2015, CITE researchers evaluated single and multi-parameter water test kits available on the market in Gujarat, India.

Among other criteria, CITE evaluated the ease-of-use of various portable water test kits among untrained users in non-lab settings. By surveying and observing end-users before and after using the product, CITE was able to gauge the perceived ease-of-use prior to using the product, the performance of users in successfully completing testing and the ability of users to successfully interpret test results. To better understand how external factors affect ease-of-use, the CITE team divided participants into two groups. Group 1 was intended to mimic a consumer's experience purchasing the product at a store; they were read aloud the instructions provided with the test kit, and brief answers were provided to their questions to simulate the support of a store clerk. Group 2 was provided a live demonstration of how to use the product, step-by-step (a process which took less than 5 minutes).

Through direct observation, the CITE team concluded that Group 2 made 62% fewer mistakes in carrying out water testing, and asked for assistance 32% fewer times as compared to Group 1. Such a blend of surveying and direct observation methodologies helped to evaluate not only the user-friendliness of a product at face value, but also how small adjustments to a product's service delivery model can affect its overall efficacy. Read more.[26]

A woman tests water quality in India. Photo by MIT CITE.

B. Conducting fieldwork

When it comes time to gather data through field-based surveying and testing, there are a number of practical and logistical items to consider. Whether your organization is directly conducting fieldwork, or you are working with or through a community partner, taking the following steps will help facilitate the field-based data collection process:

1 **Identify data collectors:** If you will not be directly administering surveys, seek out and select interviewers with experience in field research and interviews. At a minimum, interviewers

The Solar Sister team prepares for interviewing end users. Photo by MIT CITE.

should possess sufficient literacy and numeracy, and be able to effectively engage with a variety of stakeholders. If you or the interviewers do not speak the local language, make sure you have reliable interpreters with whom you are comfortable communicating.

2 **Train data collectors:** together with the interviewers and interpreters, go over your surveys and translate them if necessary. Make sure to explain why you are asking each question, and try to get input from interpreters and translators to ensure that questions are asked in an understandable and culturally appropriate fashion. You should also instruct those who will be in charge of inputting and recording data, regardless of your survey format. For more guidance, you can refer to the training protocols[27] referenced in the appendix of the Practitioner's Guide.

3 **Make a game plan:** if you haven't already, create a schedule for the time you will be conducting fieldwork, identify and map out all interviewees, and build sufficient buffer time for travel, revision of surveys and data collection tools and unforeseen delays. You will also want to set up a protocol and schedule for recording data. Some interviewers record answers as they are given, while others keep things more conversational and occasionally pause to input data. If you are using paper-based surveys, you may want to input data to a spreadsheet or other recording device the same day as you conduct the interview, while the information is still fresh in your mind.

4 **Determine compensation:** many organizations have existing protocols for compensating data collectors for their work, and interviewees for their time. Compensation could be monetary, a token of your appreciation — such as home goods or a souvenir — or nothing at all depending on your organization's policies and local customs and norms.

5 **Pilot and iterate:** if possible, data collection activities early on, with sufficient time scheduled to modify surveys and testing protocols based on the feedback and results from initial piloting. In surveys, for example, you may find that some questions are superfluous, and others need to be rephrased. For example, if you are trying to understand why someone purchased a solar lantern from a specific vendor, and you ask an open-ended question like "Why did you buy a solar lantern," they may respond with equally open-ended answers, such as "to work at night" or "because I don't have access electricity." Rephrasing the question to "Who did you purchase your solar lantern from, and why did you buy it from them rather than somewhere else," you will get more specific responses, such as "because they is where I always shop" or "because I trust the business." It is also important to avoid asking leading questions

C. Preparing technical testing protocols

Before diving into the technical-testing portion of your evaluation, there are multiple things to consider, including whether technical testing is required at all. If it is required, you should consider how to balance

resource limitations with the level of rigor required to produce satisfactory results for your evaluation and its key decision makers. These considerations include the following questions:

■ Can you afford to purchase product demos, and is it logistically feasible to acquire them for testing?

■ What facilities do you have at your disposal, and will the testing be lab-based, field-based or both?

■ If performing lab testing, can you mimic various use cases and implementation contexts?

■ How might you map technical results to your evaluation criteria and metrics, for example, will you be testing for performance, durability and other criteria?

Case Study

Solar Lantern Evaluation

In summer 2013, a team of MIT faculty and students traveled to western Uganda to conduct CITE's first-ever product evaluation: a Consumer Reports-style comparative rating of solar lanterns. Researchers conducted hundreds of surveys with consumers, suppliers, manufacturers, and nonprofits to evaluate 11 locally available solar lantern models.

Researchers from the CITE team conducted 64 semi-structured interviews with solar lantern users in Uganda. These users were asked to provide details about how they used their lanterns and which activities the products were used to illuminate. Additionally, solar-lantern users were asked which lantern characteristics were most important to them. During semi-structured interviews, CITE researchers also took note of how end users employed their solar lanterns. The data gathered in the field study allowed the CITE team to create a list of criteria, including the following subcriteria:

■ Lantern runtime

■ Time to charge

■ Task lighting capability

■ Ambient lighting

■ Brightness

■ Luminous range

■ Water resistance

While technical testing to measure these subcriteria was performed in a lab setting, much of this data can be gathered through controlled testing in many environments with simple tools, such as a lux meter (ranging from $80–$1,000 USD). For example, to test brightness, researchers measured the luminous flux per square meter that each lantern produced on the highest setting at a standard distance of 18 inches. To test water resistance, researchers developed a three-tiered testing regimen to rate lighting units based on their resistance to water. These tests involved complete submersion, exposure to heavy rain in a vulnerable orientation, and exposure to heavy rain in an "as charging" position. Representative rainfall rates were determined by examining rainfall rates from storms in various locations in the U.S., and extrapolating to Uganda. Read more.[28]

Charging a solar lantern in Uganda during a CITE evaluation. Photo by MIT CITE.

CITE researchers test water pumps in a controlled setting. Photo by Brennan Lake, Technology Exchange Lab.

- Can you perform the same testing protocols and procedures for each product?

- Will you use sensors to monitor usage patterns or performance fluctuations?

- Will you use numeric ratings? If not, how can you accurately score non-numerical ratings, e.g., "outstanding, excellent, average, poor."

- Do you or someone in your organization have the necessary technical skills to carry out the testing?

If you have decided that technical testing is required for your evaluation, you may begin to determine what type of testing to perform (controlled vs. uncontrolled) and where you will test (in the lab, field or both). By using improved cookstoves as an example, we can evaluate the tradeoffs between various types of testing. In determining the comparative performance of various cookstove products, a controlled test in a lab-based setting may help you observe discrete differences between technical aspects of different product designs, but such tests may not be representative of real-life cooking scenarios.

The Water Boiling Test,[29] published by the Global Alliance for Clean Cookstoves, offers testing protocols that may be applied to a wide variety of product alternatives by performing a universally applicable use case: boiling water. Components of the controlled water boiling test may be performed in nearly any setting, without sophisticated lab facilities. As a controlled test, you can record the amount of time it takes to boil water with each cookstove by following protocols such as these:

- Consistently start the test from the same stove temperature, whether hot or cold

- Perform the test in an environment without wind or breeze

- Implement a controlled fuel measurement procedure, before and after boiling water
- Use the same water pot, and the same volume of water for all cookstoves

Such testing could help you measure your evaluation criteria, such as cooking time and fuel efficiency. However, it may not offer insight into user-centered criteria, such as ease-of-use. In this case, adding a field-based testing component is recommended. A similar test in a field setting could involve actual end-users who are likely to use the products to boil water during meal preparation. In addition to testing cooking time, you can also interview end-users after the process is complete to determine their satisfaction levels and opinions. Protocols for such a test may include the following steps:

- Select end-users with similar experience levels relative to the cookstove (e.g., all beginners or all habitual users)
- Choose a single meal, which end-users may prepare regularly
- Measure out all ingredients and inputs equally
- Implement a controlled fuel measurement procedure, before and after making the meal
- Observe end-user cooking in a natural environment, such as their homes or community cooking spaces

While not as rigorous as a lab-based test, user-based field testing may provide a truer experience of how end-users actually engage with the products.

Throughout the testing process, it is essential to gather and record data in a way that easily maps to the criteria and metrics that you developed during the earlier stages of the evaluation process. Testing data can be logged in a number of different ways, depending on the context. This includes:

- **Offline spreadsheets:** rather than paper-based recording, digital recording through the use of spreadsheets has the added benefit of easier data analysis and visualizations. When using non-cloud based spreadsheets, such as Microsoft Excel, make sure to implement a simple version control system, especially when multiple users are inputting data into a single spreadsheet.

Helpful Tool

Sensors for remote data collection
Whether measuring the technical performance of a product, or how users interact with products over time, CITE has used data-logging sensors for both lab and field-based data collection. Developed by MIT alumnus and CITE researchers, Sensen provides a cost-effective plug-and-play solution for remote data collection. Sensen[30] sensors have been used by CITE to track usage patterns and stress points of wheelchairs; to measure temperature and moisture levels of evaporative-cooling refrigerators; and to record usage patterns and flow rates of solar-powered water pumps. In addition to the hardware component, remote sensors are often bundled with software for data analysis.

- **Online forms:** If you have a stable Internet connection at your head office, consider setting up an online form to have team members submit the data they've collected. Tools like Google Forms will automatically enter your answers into a shared spreadsheet as the team members submit each survey response.
- **Tablet entry:** Even if you didn't use tablet based data entry in the field you can later submit the responses via a tablet interface. These will then later sync via WiFi and allow your team to work from a centralized database.
- **Sensors:** Data logging through the use of sensors has the added benefit of precision data recording, while allowing for data collection over long periods of time without direct observation.

In the following section, we will explore various methods and best practices for data analysis, while also demonstrating how you can produce actual product ratings from data you gathered, allowing for easy comparison of products based on how they score on your chosen criteria and subcriteria.

Chapter 6
Analyzing data

Once you have gathered and collated data in a digital format, you may begin to translate your raw data into ratings and insights on product performance in order to drive the decision making process. Some of your data points may originate from a single, objective source. For example, through our desk study on refugee shelters we found that two models in our product list have lock features on the shelter door, while the others do not. In this case, your data points for this metric consist of a simple "yes" or "no" for each product.

Evaluating more subjective metrics, such as the number of people it takes to assemble a given shelter, may yield different results depending on the source of the data. Rather than just averaging the value of the divergent data points, you can assign weights to different data sources, in order to more appropriately reflect their accuracy. In the following example we have gathered data on the number of people required to assemble three shelters from three sources:

■ Manufacturer websites
  (Desk study)

■ Manufacturer interviews
  (1 key-informant interview)

■ Refugee camp worker interviews
  (3 key-informant interviews)

Figure 9 shows a summary of the data gathered from these three sources. The manufacturers were simply asked "How many untrained people does it take to assemble?" whereas the refugee camp workers were also asked "When was the last time you assembled this shelter?" and "How many of these shelters have you assembled?" As memory fades overtime, and efficiency and skill levels improve with repeated practice, you can incorporate this additional information in order to weigh the accuracy of their answers.

Metric G. Number of people required to assemble shelters

Metrics	Shelter #1	Shelter #2	Shelter #3
Desk Study As published on manufacturer websites:	4 people	5 people	8 people
Manufacturer interview (1 interview) "How many untrained people does it take to assemble the shelter?"	4-5 people	3-5 people	7-10 people
Refugee camp worker interviews (3 interviews to respondents A, B & C) "How many people does it take to assemble the shelter?"	A: 5-6 people B: 4-5 people C: 5 people	A: 5 people B: 5 people C: 4-5 people	A: 10 people B: 8 people C: 8-9 people
"When was the last time you assembled this shelter?"	A: 2 weeks ago B: 4 months ago C: 4 months ago	A: 4 months ago B: 2 weeks ago C: 3 months ago	A: never B: 1 week ago C: 3 months ago
"How many of these shelters have you assembled?"	A: 100 B: 50 C: 10	A: 20 B: 30 C: 100	A: 0 B: 5 C: 15

Figure 9: Interview Responses

Manufacturer

Source	Response weighting
Website	3
Interview	3

Refugee-Camp Worker

Time since last assembly	Response weighting	Previous shelters assembled	Response weighting
0-4 weeks	5	0	1
5-8 weeks	4	1-10	2
2-6 months	3	11-25	3
6+ month	2	26-50	4
never	1	50+	5

Figure 10: Weighting Responses

Shelter #1 Metric G. Number of people required to assemble shelters

Metrics	Response Value	Weighting	Weighted Response
Desk Study As published on manufacturer websites:	4 people	3	4x3=12
Manufacturer interview (1 interview) "How many untrained people does it take to assemble the shelter?"	4-5 people	3	4.5x3=13.5
Refugee camp worker interviews (3 interviews to respondents A, B & C) "When was the last time you assembled this shelter?"	A: 2 weeks ago B: 4 months ago C: 4 months ago	5 3 3	
"How many of these shelters have you assembled?"	A: 100 shelters B: 50 shelters C: 10 shelters	5 4 2	
"How many people did it take to assemble the shelter?"	A: 5-6 people B: 4-5 people C: 5 people	(5+5) = 10 (3+4) = 7 (3+2) = 5	5.5x10=55 4.5x7=31.5 5x5= 25
	Totals	28	137
	Shelter 1 score for Metric G	**137/28 = 4.9 ≈ 5 people**	

Figure 11: Calculating Weighted Responses

Criteria	Sub-criteria	Metrics	Shelter 1	Shelter 2	Shelter 3
Convenience	Set-up	• Number of people required for setup	4 people	7 people	7 people
		• Time in hours to set up shelter	6 hours	5 hours	4 hours
	Portability	• Weight	160 kg	200 kg	250 kg
		• Dimensions	187.3 ft2	204.5 ft2	269 ft2
Cost		• Cost per-unit	$1250	$700	$900
Livability	Comfort	• Number of windows	4	2	6
		• Cubic feet per person at capacity	31 ft2/person	26 ft2/person	34 ft2/person
	Safety	• Door with lock?	Yes	No	Yes
		• Penetrable with blade?	No	No	Yes

Figure 12: Translating Data into Metric Values

Based on the experience and expertise of different interview respondents and the level of trust in manufacturer information, we can assign specific weights to the different data sources. In Figure 10, we assigned the data originating from the manufacturers with a weighting of "3" since they have a deep understanding of their products, but they could also be biased. On the other hand, an experienced refugee-camp worker who has assembled over 50 shelters — including at least one in the last 4 weeks — would offer highly accurate data, to which we have assigned a rating of 10 (5 for shelter quantity plus 5 for time period).

Finally, by applying these weightings to the data we gathered through our desk study and key-informant interviews, we can develop single values for each metric. In Figure 11, we have taken the data gathered on Shelter #1 from manufacturers' websites, manufacturer interviews and refugee-camp worker interviews, and applied weightings to their responses. Specifically, we take the numeric value of each response, and multiply it by the weighting in order to produce the weighted response. We then take the weighted response sum and divide it by the sum of the assigned weightings to produce the overall metric score, which in this case is 5. As you can see, the website of the shelter manufacturer may have underreported the number of people

required for assembly (4 people), while the most experienced refugee-camp worker suggests that it realistically takes 5–6 people to set up the same shelter. By factoring in response weightings, we can get reconcile subjective divergent responses to get a more accurate estimation of the metric value. Once you are satisfied with the values assigned to each metric, you can begin to look at overall product performance across your evaluation criteria. In Figure 12 we have translated the data we gathered into metric values for three shelter products, across three of our criteria: Convenience, Cost and Livability. While looking at specific metric values across three products can help inform decision-making within isolated subcriteria (for example, Shelter 3 clearly performs best within the "Comfort" sub-criterion), it becomes increasingly complex to process multi-unit metric values across all criteria in order to determine an overall product score.

One way of resolving this complexity is by applying visual icons that indicate metric-value performance along a spectrum from "outstanding" to "poor," a method often employed by Consumer Reports, and by CITE in past evaluations. Applying these icons can be straightforward for binary yes/no metric values. For example, the "yes" value for a shelter with a door lock feature can be converted to "outstanding," while a "no" value can be converted

Ranking key

- Outstanding
- Very Good
- Average
- Marginal
- Poor

Set up (people)

- Outstanding: 1-2 ppl
- Very Good: 3-4 ppl
- Average: 5-6 ppl
- Marginal: 6-7 ppl
- Poor: 8+ ppl

Set up (time)

- Outstanding: 1-2 hrs
- Very Good: 3-4 hrs
- Average: 5-6 hrs
- Marginal: 7-8 hrs
- Poor: 9+ hrs

Weight

- Outstanding: <150 kg
- Very Good: 151-175 kg
- Average: 176-225 kg
- Marginal: 226-250 kg
- Poor: >251 kg

Dimensions

- Outstanding: <151 sq ft
- Very Good: 150-200 sq ft
- Average: 200-250 kg
- Marginal: 250-300 kg
- Poor: >300 sq ft

Space per person

(UNHCR and SPHERE minimum for indoor living space: 3.5 m2 = 38 ft2 - UNHCR[31])

- Outstanding: 58+ sq ft
- Very Good: 48-57 sq ft
- Average: 38-47 sq ft
- Marginal: 28-37 sq ft
- Poor: <28 sq ft

Windows

- Outstanding: 5+
- Very Good: 4
- Average: 3
- Marginal: 2
- Poor: 1

Cost

- Outstanding: 600-699 USD
- Very Good: 700-799 USD
- Average: 800-899 USD
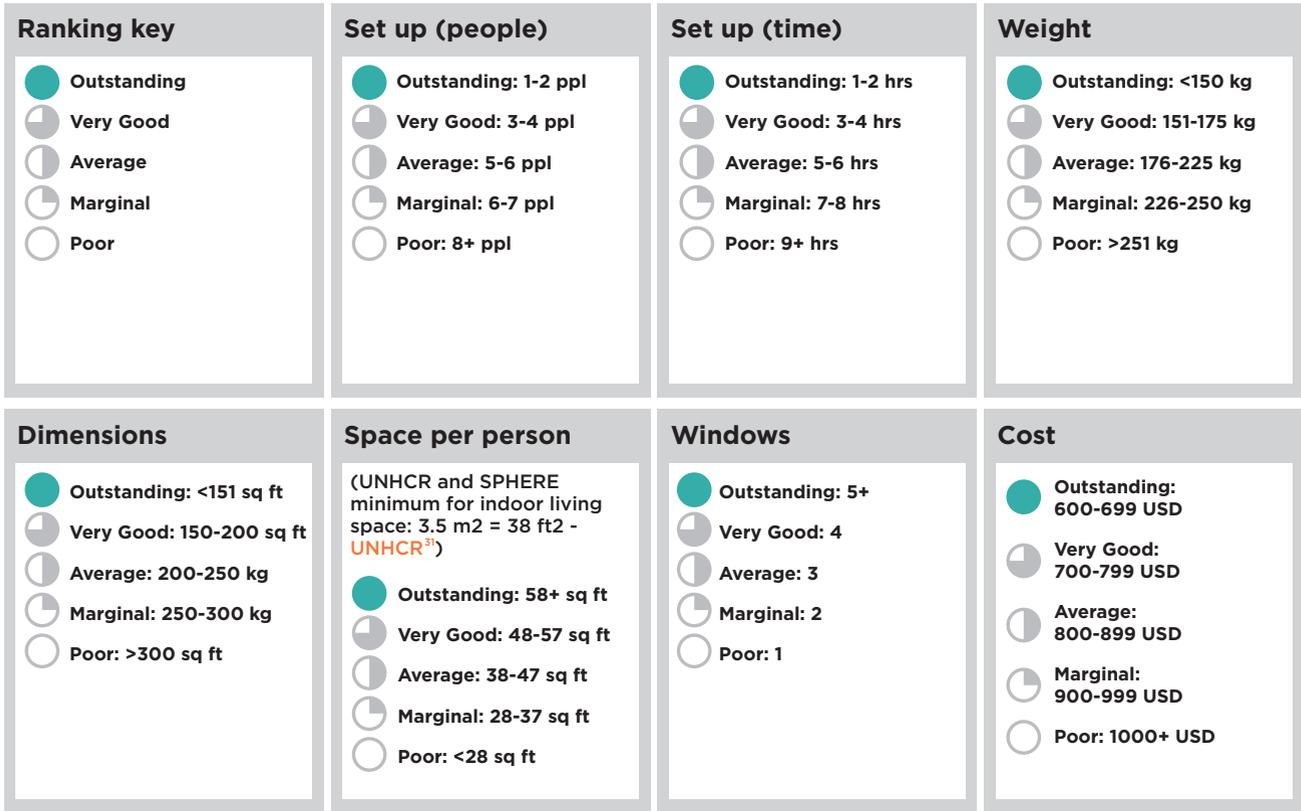- Marginal: 900-999 USD
- Poor: 1000+ USD

Figure 13: Ranking Metric Values with Visual Icons

to "poor." Ranking other metric values may be more subjective, such as cost, which largely depends on the ability and willingness to pay of decision makers. In order to rank subjective metric values, it is necessary to define numeric ranges that correspond to performance. In the following table, we have defined numeric ranges for each metric value with its corresponding performance ranking. In defining these ranges, it is crucial to work closely with key decision makers and stakeholders, in order to get their consensus on what constitutes outstanding, very good, average, marginal and poor rankings for each metric value.

Once you have set these ranges, you may simply apply icon rankings to the corresponding metric values, in order to produce full comparative rankings that can be visualized across multiple criteria, as detailed in Figure 14. Now that we are able to more easily compare products across different criteria, how can we go about determining how products rank compared to one another? The answer to this question is that there is no single correct way

to calculate overall product scores and it may not be necessary to produce an overall product score; however, there are a number of approaches you can adopt to most closely align your scoring system with your priorities as a decision-maker. These approaches include calculating overall product scores by:

- Averaging criteria ranking
- Averaging metric ranking
- Applying weightings to criteria in accordance to their importance and averaging weighted criteria

In Figure 15, we have applied visual rankings for our products across all seven of our criteria. At face value, we can calculate overall product scores by averaging metric rankings. No matter how we calculate overall product score, we assign universal values to each metric ranking as follows:

Criteria	Sub-criteria	Metrics	Shelter 1	Shelter 2	Shelter 3
Convenience	Set-up	• Number of people required for setup			
		• Time in hours to set up shelter			
	Portability	• Weight			
		• Dimensions			
Cost		• Cost per-unit			
Livability	Comfort	• Number of windows			
		• Square feet per person at capacity			
	Safety	• Door with lock?			
		• Penetrable with blade?			

Figure 14: Visualization Product Scores by Metrics

Outstanding = 5/5

Very good = 4/5

Average = 3/5

Marginal = 2/5

Poor = 1/5

In this scenario, averaging across unweighted metric values gives overall product scores with shelter #1 emerging as the clear winner, as calculated and demonstrated in column two of the table in Figure 16. However, we can also take a more nuanced approach that takes into consideration the priorities of multiple decision-makers and stakeholders within your program. For example, if your organization or decision-makers put a maxim on budget, you can apply a greater weighting to the cost criterion. By assigning a 50% weight to the cost criterion in the example below, we see that Shelter #2 is given a higher score.

The same method may be applied to other criteria at your discretion. If your greatest priority lies with the comfort and well-being of beneficiaries, you can apply a greater weighting to the Livability criterion. If you are more concerned with the efficiency and satisfaction of refugee-camp workers, you may then decide to weigh Convenience greater than other criteria. In both cases, Shelter #1 again receives the highest score.

Criteria	Sub-criteria	Metrics	Shelter 1	Shelter 2	Shelter 3
Cost		• Cost per-unit	○	◔	◔
Livability	Comfort	• Number of windows	◔	◔	●
		• Square feet per person at capacity	◖	◖	◔
	Safety	• Door with lock?	●	○	●
		• Penetrable with blade?	●	●	○
Convenience	Set-up	• Number of people required for setup	◔	◔	◔
		• Time in hours to set up shelter	◖	◖	◔
	Portability	• Weight	◔	◖	◔
		• Dimensions	◔	◖	◔
Durability	Wind resistance	• Wind speed to failure	◖	◔	◔
	Water resistance	• Leaks with high pressure hose (y/n)	●	●	○
Availability	Shipping time	• Time in days to ship from distributor to site	◔	◖	◔
	Production capacity	• Number of shippable SKUs per month at peak production	◔	◔	◖
Customization		• Number of layout options	◖	◖	◖
		• Shelves/Hooks (y/n)	●	○	●
Compliance		• Meets local regulations (y/n)	●	●	○

Figure 15: Full Visualization of Product Scores by Metrics

In considering this particular example, it is worthy to note that your evaluation results can very well provide data-based evidence for influencing the opinions of key decision-makers. While you may have budget-conscious decision makers who are inclined to make a procurement decision in favor of Shelter #2, you can provide them with data that suggests that spending more per-unit on shelters could ultimately lead to better outcomes in terms of beneficiary well-being and human resource efficiency on the ground. Such insights could perhaps even persuade your program's funders to increase financial support based on this data-driven rationale.

This is just one approach toward data analysis and visualization. For information on other data analysis tools, be sure to check out the Further Reading & Additional Resources section at the end of the *Practitioner's Guide*.

Metrics	Shelter 1	Shelter 2	Shelter 3
Unweighted	0.75	0.61	0.58
Cost 50%	0.49	0.7	0.49
Livability 50%	0.78	0.59	0.63
Convenience 50%	0.75	0.59	0.55

Figure 16: Weighting Overall Product Score by Criteria

Data-logging sensor attached to a vegetable cooling device. Photo by Lauren McKown, MIT D-Lab.

Chapter 7
Reporting results

At this point you have conducted your comparative evaluation, and hopefully gained helpful insights on how different products will perform within the context at hand, so what next? In translating insights into effective action, it is first necessary to determine whom specifically the evaluation is for, and who the key decision-makers are.

If you are the ultimate decision-maker, great! You can now take steps toward procuring and implementing the products that best respond to the specific needs, challenges and opportunities within your context. For additional guidance on developing appropriate distribution models and business modeling, you can use many of the tools covered in the Scoping Study section — such as stakeholder mapping, and user personas. For references on monitoring and evaluation, be sure to check out the Further Reading & Additional Resources section.

If you are conducting this evaluation in order to inform other key decision makers, whether within your organization or externally, it is necessary to create an evaluation report in order to convey your findings in a comprehensive and digestible format. Indeed, a report may be required as a deliverable or for reasons of compliance with a donor or funding agency. In this case, it is important to present your findings in a way that satisfies such requirements, while also making it clear for readers as to how certain decisions may lead to different outcomes. Clearly articulating your evaluation methodology and approach also gives decision makers confidence in the findings.

In the previous section, we demonstrated how weighing organizational priorities and preferences differently can lead to a variety of outcomes vis-a-vis product scoring. In your report, it can be helpful to go beyond simple product scores and point out trade offs between products, while also illustrating how alternative decisions can lead to different outcomes through hypothetical use-case scenarios. As part of CITE's evaluation of solar-powered water pumps[32] in India, CITE researchers developed a System Dynamics model of the effects of various solar-pump implementation policies on the food-water-energy nexus in India. In this case, such modeling can help policy makers map out different scenarios and solar-pump implementation schemes in order to simulate outcomes.

Regardless of who your intended audience is, consider documenting your findings for future use by others. If your evaluation can help inform future decision-making without duplication of efforts, then it is worth keeping track of not only your evaluation results, but also the steps and processes you employed to achieve them. In doing so, be sure to point out caveats on the generalizability of your findings. For example, CITE's evaluation on solar lanterns was based on context-specific criteria that responded to user needs and preferences that were germane to Uganda. In this case, the final evaluation report made clear that product criteria were developed around the Ugandan context, even if scores around specific metrics are generalizable regardless of geographic context.

It is also important to note that in addition to geographic context, over time product characteristics and their relevance to performance standards may change. The emergence of new products and models, new supply chains, price fluctuations, and changing standards of living will all affect the relevance of your evaluation findings over time. In this case, it is helpful to document your evaluation process so that it may be more easily updated and modified without having to reproduce the entire evaluation.

Summary/Conclusion

We hope you have enjoyed reading the Practitioner's Guide, and acquainting yourself with the multi-faceted process of evaluating products and technologies for global development. Our goal is to provide you with a concise and actionable roadmap

to assist you in performing comparative evaluations for new or existing programs.

While our framework is designed to be practical and replicable, it is also part of an ever-evolving methodology toward improving the quality and efficiency of the evaluation process. As a development-oriented research organization, CITE's goal is to develop a deep understanding of what makes products successful in emerging markets. Our evaluations provide evidence for data-driven decision-making by development workers, donors, manufacturers, suppliers, and consumers themselves. In addition, CITE evaluations lead to significant development insights, helping us better understand and address global development challenges. Beyond evaluations, CITE seeks to share its approach through global replication, education, and dissemination.

If you are interested in learning more about CITE, The Technology Exchange Lab, and partnership opportunities, we invite you to contact the following focal points for more information on how to collaborate:

Joanne Mathias
CITE Associate Director
mathiasj@mit.edu

Brennan Lake
Technology Exchange Lab Programs Director:
brennan@techxlab.org

Further reading & additional resources

For additional resources and tools related to identifying, evaluating and implementing global development solutions, check out the references below.

Identifying challenges

1. UN Office for the Coordination of Humanitarian Affairs "Multi-Cluster/Sector Initial Rapid Assessment (MIRA)" Inter-Agency Standing Committee (IASC). Geneva. 2012. Link-phrase: Humanitarian Crisis Response https://docs. unocha.org/sites/dms/documents/mira_final_ version2012.pdf

2. Verploegen, E. "D-Lab's Energy Assessment Toolkit" MIT D-Lab. Cambridge, MA. April 2017. Link-phrase: Community-based energy needs. https://d-lab.mit.edu/off-grid-energy/roadmap/ assessment-toolkit

3. UCLA Center for Health Policy Research "Key Informant Interviews" Health DATA Program – Data, Advocacy and Technical Assistance. Los Angeles. Link-phrase: Interview guide http:// healthpolicy.ucla.edu/programs/health-data/ trainings/Documents/tw_cba23.pdf

4. Keast, G. "Water in Schools Monitoring Package" UNICEF. April 2011. Link-phrase: Checklist https://www.unicef.org/wash/schools/files/ wash_in_schools_monitoringpackage_.pdf

5. D-Lab. "Principles for conducting human-centered field research" Lean Research MIT D-Lab. Accessed November 2017 Link-phrase: Lean research https://d-lab.mit.edu/lean-research

6. mSurvey. Link-phrase: mSurvey http://msurvey. co.ke/

7. CommCare. Accessed 11-20-2017 Link-phrase: CommCare https://www.commcarehq.org/ home/

8. Frontline. Accessed 11-20-2017 Link-phrase: Frontline http://www.frontlinesms.com/

9. Kobo. Accessed 11-20-2017 Link-phrase: Kobo http://www.kobotoolbox.org/

10. Magpi. Accessed 11-20-2017 Link-phrase: Magpi https://home.magpi.com/

11. Formhub. Accessed 11-20-2017 Link-phrase: Formhub http://formhub.redcross.org/

12. DataWinners. Accessed 11-20-2017 Link-phrase: DataWinners http://hni.org/what-we-do/data-collection/datawinners/

13. CITE. "Vegetable Cooling and Storage Evaluation" MIT CITE. Accessed November 2017. Link-phrase: Comparative evaluation http:// cite.mit.edu/vegetable-cooling-and-storage-evaluation#

Selecting products

14. Technology Exchange Lab. Accessed 11-20-2017 Link-phrase: Technology Exchange Lab https:// www.techxlab.org/

15. USAID Global Innovation Exchange. Accessed 11-20-2017 Link-phrase: Global Innovation Exchange https://www. globalinnovationexchange.org/

16. Global Alliance for Clean Cookstoves. Accessed 11-20-2017 Link-phrase: Global Alliance for Clean Cookstoves http://cleancookstoves.org/

17. Lighting Global. Accessed 11-20-2017 Link-phrase: Lighting Globa: https://www. lightingglobal.org/

18. CITE. "Wheelchair Evaluation" MIT CITE. Accessed November 2017. Link-phrase: Comparative evaluation of various wheelchair models http://cite.mit.edu/wheelchair-evaluation

Developing evaluation criteria

19. Corinne Carland, Gilberto Montibeller, Jarrod Goentzel. "Evaluating Business Criteria for Scaling Stock of Malaria Rapid Diagnostics" MIT CITE. Accessed November 2017. Link-phrase: MIT researchers evaluated the private sector uptake for mRDTs in Uganda http://cite.mit.edu/reports/ malaria-rapid-diagnostic-test-evaluation

Conducting scoping studies

20. CITE. "Technology Evaluation for Global Development" MIT CITE. Accessed 11-20-2017 Link-phrase: Online Course https://www.edx. org/course/technology-evaluation-global-development-mitx-cite101x#!

21 Jarrod Goentzel et al. "Feed the Future Uganda: Market Systems Monitoring Activity" MIT CITE. Accessed November 2017. http://humanitarian.mit.edu/projects/feed-the-future-uganda

22. Jarrod Goentzel, Courtney Blair, Erica Gralla "New Approaches to Measuring Market Dynamics in Uganda" MIT CITE. July 2017. https://agrilinks.org/post/new-approaches-measuring-market-dynamics-uganda

23. Pamela Vaughan "How to Create Detailed Buyer Personas for Your Business" Hubspot. Accessed 11-20-2017 Link-phrase: Free template https://blog.hubspot.com/marketing/buyer-persona-research

24. IDEO "Design Kit.". IDEO. Accessed 11-20-2017 Link-phrase: IDEO's free Design Kit http://www.designkit.org/

25. Rebecca Spohrer. "How to Drive Adoption of Social Innovations Amid Social Taboos" Acumen. November 2015. Accessed 11-20-2017 Link-phrase: Acumen's guide https://www.plusacumen.org/journal/how-drive-adoption-social-innovations-amid-social-taboos

26. Joan Churchill, Eric von Hippel, Mary Sonnack "Lead User Project Handbook: A practical guide for lead user project teams" October 2009. Link-phrase: Lead user methodology https://evhippel.mit.edu/teaching/

Collecting data & testing products

27. Jennifer Green, Susan Murcott et al. "Water Test Kit Evaluation" MIT CITE. Accessed November 2017. Link-phrase: CITE researchers evaluated single and multi-parameter water test kits: http://cite.mit.edu/reports/water-test-kit-evaluation

28. Derek Brine, Daniel Frey, Jennifer Green, Jarrod Goentzel et al. "Solar Lantern Evaluation" MIT CITE. Accessed November 2017. Link-phrase: CITE researchers evaluated solar lanterns: http://cite.mit.edu/reports/solar-lantern-evaluation

29. Global Alliance of Clean Cookstoves. "The Water Boiling Test (Version 4.2.3)" March 2014. Accessed 11-20-2017 Link-phrase: Water Boiling Test https://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/399-1.pdf

30. Sensen. Accessed 11-20-2017 Link-phrase: Sensen: http://www.sensen.co/

Analyzing data

31. UNHCR "Camp Planning Standards (planned settlements)" Accessed 11-20-2017. Link-phrase: UNHCR https://emergency.unhcr.org/entry/45582/camp-planning-standards-planned-settlements

Reporting results

32. Jennifer Green et al. "Solar Water Pumps: Technical, Systems, and Business Model Approaches to Evaluation" MIT CITE. November 2017. Accessed November 2017. Link-phrase: CITE's evaluation of solar-powered water pumps http://cite.mit.edu/reports/solar-powered-water-pumps

Acknowledgments

Authors

Brennan Lake
Programs Director, Technology Exchange Lab

Éadaoin Ilten
Operations Manager, Technology Exchange Lab

Jennifer Green
Sustainability Evaluation Lead, CITE

Special thanks to the following team members from CITE and MIT:

- Joanne Mathias
- Professor Daniel Frey
- Professor Bishwapriya Sanyal
- Kendra Leith
- Lauren McKown
- Victor Park
- Jonars Spielberg
- Jarrod Goentzel
- Mark Brennan
- Sara Pesek
- Amit Gandhi
- Eric Verploegen