



Behavior Change Interventions in Practice:

A synthesis of criteria, approaches, case studies & indicators

Authors:

Katie Williamson

Senior Associate, Center for Behavior & the Environment, Rare

Philippe M. Bujold

Senior Associate, Center for Behavior & the Environment, Rare

Erik Thulin

Behavioral Science Lead, Center for Behavior & the Environment, Rare

Recommended citation:

Williamson, K., Bujold, P. M., & Thulin, E. (2020). Behavior Change Interventions in Practice: A synthesis of criteria, approaches, case studies & indicators. Rare Center for Behavior & the Environment and the Scientific and Technical Advisory Panel to the Global Environment Facility.

Acknowledgments:

We would like to acknowledge the valuable review and comments on this report from Edward Carr (STAP), Graciela Metternicht (STAP), Mark Stafford Smith (STAP), Guadalupe Duron (STAP secretariat), Christopher Whaley (STAP secretariat), and Kevin Green (Rare BE.Center); the support of Andrea Wilk (Rare BE.Center) and Camille Freeman (Rare BE.Center) in case study development; research by Milan Urbanik (London School of Economics) and Ganga Shreedhar (London School of Economics) identifying behavior change frameworks and case studies; and Corinn Weiler (Rare) and Kyla Timberlake (Rare) for graphics development and document design. Cover photo by George Stoyale.

This report was commissioned and funded by the Scientific and Technical Advisory Panel to the Global Environment Facility.



This work is licensed under CC BY 4.0.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0>

Table of Contents

| | |
|--|-----------|
| Introduction: Behavior Change Interventions in Practice | 4 |
| Glossary | 5 |
| Summary: Criteria for Behavior Change Interventions | 8 |
| Criteria for Behavioral Context Analysis | 10 |
| Criteria for Intervention Development | 12 |
| Criteria for Monitoring and Evaluation | 14 |
| Part I: Review of Behavior Change Approaches | 17 |
| Introduction | 18 |
| Observations and Recommendations For Selecting a Behavior Change Approach | 20 |
| Review of Selected Behavior Change Approaches | 24 |
| Appendix A: Complete List of Reviewed Organizations and Their Behavior Change Approaches | 39 |
| References | 40 |
| Part II: Case studies | 42 |
| Case Study 1: Water Consumption in Costa Rica | 44 |
| Case Study 2: Water Consumption in Colombia | 47 |
| Case Study 3: Increasing Farm Productivity in Kenya | 50 |
| Case Study 4: Wild Meat Consumption in Brazil | 53 |
| Case Study 5: Sustainable Farming in Mexico | 55 |
| Case Study 6: Reducing Overfishing in Indonesia | 58 |
| Case Study 7: Silvopasture in Colombia | 61 |
| Case Study 8: Increasing Rhino Monitoring in Namibia | 64 |
| Conclusion: Case Study Recommendations | 67 |
| Part III: Selecting, Measuring, Monitoring, and Evaluating Behavior Change Indicators | 70 |
| Indicator Selection | 72 |
| Indicator Measurement | 76 |
| Framework for Indicator Monitoring and Adaptive Management | 80 |
| Frameworks for Evaluating Changes in Indicators | 85 |
| Cross-Context Generalizability | 89 |
| Conclusion: Recommendations On Indicators For Behavior Change Programming | 90 |
| References | 92 |

Introduction: Behavior Change Interventions in Practice

This document serves as a complement to the literature review on behavior change interventions specific to the environmental sector. Rather than presenting evidence on what interventions have been successful or unsuccessful, the parts of this document cover the how of behavioral design. We first outline a summary of criteria for behavior change interventions that draws from the insights and best practices presented in the three Parts to follow. These important criteria serve as a tool for current and future behavior change designers and environmentalists. Part I covers behavior change approaches, including criteria and assessments of current behavior change design methodologies and tools. Part II provides eight, in-depth case studies of behavior change interventions, covering different environmental topics and geographies, to give additional detail as to what the stepwise process of designing behavior change solutions looks like. Part III describes how to select, measure, monitor, and evaluate indicators of behavior change so that the solution is supported by a robust theory of change and evaluation protocol. Together, these provide a rich set of analytic and descriptive guidelines to achieve more effective, sustainable, and responsible behavior change work.

Glossary

Actors: People whose behavior directly or indirectly affects program outcomes

Adaptive management: a process of updating and improving how a program is managed based on data and feedback about what is working or not working

Attitude: An evaluation of something, ranging from negative to positive

Barriers: Forces, attitudes, beliefs, or other reasons that prevent someone from doing a behavior

Behavior: An action that a person takes in response to something (a stimuli)

Behavior-actor pair: A grouping that identifies a behavior and who is doing it

Behavioral insights: The findings that result from analyzing patterns in how people tend to behave.

Behavioral system: A network of actors, mapping how each actor's behavior influences each other's capacity to act and their interest in doing so

Behavior change approach: A methodology for changing behavior, often drawing upon principles of behavioral design

Behavioral indicator: A measurement that signifies behavior (or something that approximates it) has changed

Behavior change intervention, programming: A set or sequence of activities that aim to influence actors to adopt target behaviors to achieve a certain outcome

Behavioral design: An approach that blends insights from the design and behavioral and social science fields

Belief: Something that someone accepts to be true

Bias/Cognitive bias: A way of thinking that systematically deviates from rational choice

Control: A level of an independent variable a person or group is assigned to in a study that receives no additional intervention

Counterfactual: A comparison for an intervention to assess its impact that shows what would have happened if no intervention had taken place

Cross-context generalizability: The degree to which something applies to other socio-ecological contexts

Design thinking: A creative and iterative process for developing, designing, and testing innovative solutions, often used in combination with human-centered design

Difference-in-difference: A quasi-experimental method that compares the pre-post change (difference) in outcomes for the treatment group with the change in outcomes of a comparison group

Direct observation: Type of behavior measurement based on directly observing behavior, rather than using a proxy or self-report measurement

Disaggregation: A data reporting process that shows how an intervention may have impacted different groups differently

Doer/non-doer analysis: A comparison of the motivations and barriers for people who are already doing the target behavior and those not doing the target behavior

Durability: The degree to which an intervention's effects persist during an intervention period and after the intervention has ended

Dynamic programming: Making live programmatic decisions about phase transitions, expansion, or termination based on real-time monitoring of psychological and social states of the target actors

Human-centered design: An approach or mindset to problem-solving that centers people's needs and goals in solution designs, often combined with design thinking

Matching: A quasi-experimental method that builds a comparison group by identifying units that are similar to each of the treatment units based on a relative set of observable characteristics

Motivations: Forces, attitudes, beliefs, or other reasons that encourage someone to do a behavior

Outcomes: The behavioral, social, or other goals or objectives a program is trying to achieve

Outputs: The components of a program that help to show how it achieved its outcomes and may serve as intermediary objectives

Pre-post comparison: A study where a treatment effect is estimated by subtracting the base-line value from the value after treatment

Program activities: The parts of an intervention that are implemented to change behavior, such as training sessions, pledges, incentive mechanisms, etc.

Prototype: A small-scale version of a behavioral solution that captures its essential features and can be tested with target actors

Proxy measures: Type of behavior measurement that uses outcomes assumed to be tightly related to the target behavior

Psychological indicator: A measurement that signifies a belief, attitude, or preference (or something that approximates it) has changed

Psycho-social state: Beliefs, values, expectations, and social relations that result from program activities, and other psycho-social states and also influence future behavior

Psycho-social theory of change: A theory of change that links intervention components to psychological or social changes, leading to behavioral outputs and environmental and social out-comes

Pulse monitoring: Assessing key psycho-social indicators on a frequent basis throughout pro-gram delivery

Quasi-experimental methods: Evaluation methods that infer the causal effect of an intervention without randomization when assigning individuals to treatment conditions

Randomized evaluations, Randomized Control Trials (RCTs): Evaluation methods where individuals are randomly assigned to treatment conditions

Self-report measures: Type of behavior measurement where the rate or intensity of a behavior is inferred through responses from instruments such as surveys

Social indicator: A measurement that signifies a social state, structure, or factor (or something that approximates it) has changed

Social marketing: The application of techniques from marketing to shift behavior to benefit individuals and society

Socio-ecological system: A system of interdependent linkages between ecological factors, social and cultural factors, and institutions at different scales that continually adapt over time

Stakeholders: Individuals or groups who have an interest in environmental outcomes or will be affected by a project and program

Study condition: A level of an independent variable a person or group is assigned to in a study

Study treatment: The intervention an individual or a group receives, based on the condition to which they were assigned

Systems thinking: An approach that synthesizes how parts of a system relate to, influence, and cause one another, often through feedback loops

SUMMARY

Criteria for Behavior Change Interventions

Introduction

Through our review of case studies, design frameworks, and indicator development for behavior change in this report and overall best practices in behavior change design, we identified a series of criteria for behavior change programming. These were selected to identify program design features that achieve effective, sustained behavior change in the environmental space. While many of these criteria do involve the incorporation of social and behavioral science, they are also strongly informed by the practical experience of designers for creating tailored behavior change solutions.

These criteria fall into three broad categories:

Behavioral context analysis criteria refer to establishing a clear understanding of the overall socio-ecological system that behavior is embedded in, a clear definition of the target behaviors and actors, and the motivations and barriers of those target actors.

Intervention development criteria refer to building insights developed in that behavioral context analysis into a socially and psychologically informed solution likely to achieve effective behavior change.

Monitoring and evaluation criteria include creating a framework that allows for the generation of applicable and generalizable data for decision making that moves beyond simply measuring program inputs and outcomes to the psychological and social states of the actors involved.

These criteria are summarized as follows:

- Identify clear and specific behavior-actor pairs
- Conduct contextual analysis to understand the socio-ecological system in which the actor and behavior is embedded
- Conduct analysis of the target actors' needs and experiences to identify motivations and barriers
- Engage a range of relevant stakeholders throughout different scales of roles
- Base the design of the intervention on a psycho-social theory of change, linking intervention components to psychological or social changes, leading to behavioral outputs and environmental and social outcomes
- Incorporate the motivations and barriers of a diverse set of actors at different scales sufficient for addressing the problem
- Rather than relying on a single behavioral insight, draw on multiple psychological and social principles to address the various barriers and motivations actors throughout the system may have
- Throughout the intervention development process, encourage the development of intervention prototypes, and test those prototypes with the target actors prior to scaled implementation
- Monitor and evaluate the intervention not just against environmental outcomes, but also behavioral, psychological, and social outputs
- Evaluate the program against a valid counterfactual to determine the causal effect of the intervention
- Continue evaluation beyond the cessation of the program to ensure that the environmental outcomes, as well as the psychological and social outputs, are sufficiently durable

Criteria

Criteria for Behavioral Context Analysis

Developing an understanding of what behaviors need to be taken by which actors, what socio-ecological context those actors inhabit, and what barriers and motivations they may have.

Identify clear and specific behavior-actor pairs

Programs that aim to change behavior need to identify the precise actions they wish to be taken, and the specific target actors they wish to take that action. These are known as behavior-actor pairs. This is distinct from programs that may aim to, for example, change attitudes or beliefs without a clear hypothesis as to why changing those attitudes or beliefs will result in the change of a target behavior. It can also be contrasted with programs that aim to change an amorphous set of behaviors, such as “encouraging climate-smart agriculture,” without a clear set of specific actions.

Being specific about which actors designers aim to influence is also critical. For example, a campaign might aim to increase the use of composting in a farming community. While this campaign may be occurring in a single community, it is likely not trying to change the behavior of all community members. In fact, the campaign may only be targeting a particular subset of community members, such as farmers. This is not to say that other members are not important for achieving this behavioral outcome, but it explicitly recognizes that their role, and therefore the target behavioral outcome, is different. For example, the children of farmers might be significant influences on their parents. Therefore, children might be target actors for the behavior of telling their parents how important it is to compost, despite them not being the direct actor themselves. In this way, it is important to identify not only the direct behavior-actor pair that is closest to the target outcome, but also all the indirect pairs that create an enabling environment throughout the social system.

Identifying behavior-actor pairs allows for precisely targeted programming. These pairs should serve as the target actors and behaviors for any intervention component, as well as guiding whether a particular component should be included in an intervention or should be revised. Only components that are expected to drive an actor to adopt their target behavior should be incorporated. For examples of where behavior-actor pairs facilitated better programming, see the Brazil and Indonesia case studies in Part II, as well as the Part I for information on how different design frameworks approach this from a practical standpoint.

Conduct contextual analysis to understand the socio-ecological system in which the actor and behavior is embedded

The actors we aim to influence do not exist in a vacuum. Their ability, as well as their interest in taking a target action, is not determined simply by their own mental processes: it is just as significantly determined by the wider cultural context that enables and limits these actors' actions. An understanding of this wider cultural context goes beyond generalities of how a group acts or what they believe. It involves identifying the systems of power, institutions, and structural forces that shape actors' identities and social roles. This has implications for constructions of gender, race and ethnicity, socioeconomic status (SES), and religious affiliation—and that, in a given context.

The program designer will need to determine which actor's identities are, in fact, relevant in enabling or limiting the actors from performing their target action. It is important not to essentialize individuals as being a part of one group or another. In fact, very few of these identities are mutually exclusive, meaning that individuals exist as an intersection of those differing dimensions. Those intersections may present different abilities and limitations than the identities themselves. For example, while women and all ethnic groups, independently, may be allowed to participate in a meeting, women of a particular ethnicity (i.e., the intersection of gender and ethnicity) may be excluded.

Understanding the socio-ecological system also involves careful examination of the relational dynamics between the implementer of an intervention and the various actors in the behavioral system. For example, while participating in an intervention may be perceived by the implementer as voluntary, the actor may perceive the implementer as in a relative position of power. This can make compliance feel required from the perspective of the actor. It is similarly critical to understand the trust dynamics that are present both within a community of actors as well as between that community and program implementers.

It is not possible to identify an exhaustive list of all of the elements of a socio-ecological system that may be relevant for any intervention. Rather, the relevant factors are intervention and context-dependent. Program developers should employ social science data collection tools, including interviews, surveys, focus groups, and observation, to determine what social factors may limit or enable actors. These socio-ecological factors will be critical to intervention development, both to understand how to best promote action given that unique system as a constraint, as well as leverage elements of that system to promote program effectiveness.

Some additional tools for understanding the behavioral system are from the field of systems thinking. Concept modeling and systems mapping are tools that help designers to draw the links between elements in a system and labeling the enabling, reinforcing, and restricting forces among them. Stakeholder mapping focuses specifically on the relationships between actors in the same context relative to a behavior. These tools are useful in plotting important actors, behaviors, and contextual factors all in one place as a reference for creating later interventions.

Conduct analysis of the target actors' goals and experiences to identify motivations and barriers

Beyond elements identified in the socio-ecological analysis, understanding the experiences of the target actors, as well as their self-identified—rather than designer assumed—needs, allows for a program designer to identify motivations and barriers for an actor to adopt a target behavior. Starting with an analysis of past experiences can reveal why an actor might have tried, whether successfully or not, to engage in the target behavior. This can demonstrate both the needs that led them to attempt the behavior and highlight the barriers that prevented them from achieving that goal.

Analyzing their past experience can reveal what might be effective for motivating an individual to act, and what barriers may prevent them from doing so. In a context where some people are engaging in the target behavior and some are not, it can be useful to compare what differentiates those 'doers' and 'non-doers'. This analysis can reveal possible socio-structural limitations on actions such as discussed above, as well as the individual-level motivations or barriers leading some to act while others did not. The intervention designer can then use this intervention to most effectively promote the target behavior in the target actor's own terms, rather than those assumed by an outside group. Motivations and barriers can often be identified through surveys, interviews, behavioral journey mapping, and other data collection techniques. Case studies in Indonesia and Namibia in Part II of this report present strong examples of the value of identifying these motivations and barriers, while Part I gives additional information on how different behavioral change frameworks approach collecting this data.

Engage a range of relevant stakeholders at different scales of the behavioral system

Identifying and engaging relevant stakeholders is important for both an effectively designed intervention as well as building stakeholder support. Identifying what makes a stakeholder relevant is context-dependent. However, it is important to note that relevance does not simply mean those in positions of power or authority. In fact, those in positions of power often have both interests and beliefs that differ from those who are most affected by an intervention. Relevance, therefore, entails representation from across the behavioral system, with a particular emphasis on actors represented in the various direct and indirect actor-behavior pairs designers previously identified. Similarly, interventions often provide costs and benefits unequally; it is critical to incorporate not only those who might benefit from an intervention, but also those who bear its costs. This means ensuring that the various identities who may interact with the intervention are represented. This could include representation across

gender, ethnicity, and SES, but also a host of other social differences identified in implementers' analysis of the socio-ecological system.

It is important to note that engaging stakeholders does not simply mean cursory consultation or the provision of information. Instead, engagement should be seen as a consistent and sustained dialogue—as opposed to simply presenting a pre-established intervention to stakeholders for approval. Stakeholder dialogue should instead be sustained throughout the intervention design process so that stakeholders' perspectives and insights are incorporated into an understanding of the behavioral and social contexts during the intervention's development.

Criteria for Intervention Development

Building an effective behavior change intervention which is firmly grounded in the insights developed in the behavioral context mapping to specifically target the barriers and motivations of target actors with psychologically and socially informed strategies

Base the design of the intervention on a psycho-social theory of change: link intervention components to psychological or social changes, behavioral outputs, and environmental and social outcomes

In order to develop an effective yet streamlined intervention, it is important for that intervention to focus specifically on those changes necessary to bring about the desired behavioral shift. Framing the intervention as part of a cohesive Psycho-Social Theory of Change (PS-ToC) allows for a highly precise intervention, focusing only on those elements that are required to bring about a change in behavior.

A ToC traditionally expresses the logic of an intervention by showing which intervention components are expected to achieve what intermediate outputs, and how those intermediate outputs are expected to affect program outcomes. However, these intervention components almost never directly lead to an intermediate output, like a behavior. Instead, they work by changing the psychological or social state of the actor, which then allows for, or motivates, that actor to adopt the behavior. These changes to psychological and social states are the critical glue linking together intervention components and behavioral outputs, but they are critically lacking from traditional ToCs. A PS-ToC draws on the findings of preliminary behavioral context mapping, and includes the explicit representation of the ways in which any given component of an intervention is expected to change the psychological or social state of a target actor (or actors)—this should then lead to an intermediate behavioral outcome. A PS-ToC also includes the explicit acknowledgment of the enabling social context that is (or will be) necessary for an intervention to be effective.

Building a behavior change program out of a PS-ToC presents a number of key benefits over a traditional ToC. First, it ensures that all intervention components are explicitly linked to insights gleaned from the behavioral context analysis rather than relying on the intuitions of a program designer. Second, it allows for each component of the intervention to be objectively evaluated relative its efficacy in changing the target psychological or social state(s), and to identify any extraneous elements in an intervention that would have no effect (direct or otherwise) on the aforementioned social and psychological states.

For more details on the components of a PS-ToC and an example from increasing reserve compliance in coastal fisheries, please refer to the Psycho-Social Theory of Change section of Part III of this report.

Incorporate the motivations and barriers of a diverse set of actors at different scales sufficient for addressing the problem

In the framing of the behavioral context, designers will likely have identified a variety of actor-behavior pairs through techniques such as systems mapping, interviews, and other qualitative data gathering techniques. And looking at each of those actor-behavior pairs, designers will have identified the various identities held by the actors that correlate with different motivations and capacities. Now that designers are developing the intervention, it is critical that they incorporate a nuanced understanding of various identities and social positions. This avoids the risk of designing an intervention fit only for the motivations and capacities within the socio-ecological system of those most prominent. Instead, an effective intervention is one designed to address the diverse needs of different actors with different identities, doing so for all the relevant actor-behavior pair.

Similarly, it is important for an intervention to address the various actor-behavior pairs across various scales, rather than focusing solely on those actors whose actions directly contribute to the program outcome. It is just as critical to ensure that programming targets those actor-behavior pairs that create the necessary enabling conditions that motivate actors to take action. One can aid in the creation of an enabling environment by providing material support or by building social capacity; think someone in a position of power or prestige allowing and encouraging someone of a different social standing to take action. The case study in Namibia presents a clear example of design that takes account of the diverse needs of different identities within a given actor-behavior pair. The case studies in Indonesia, Mexico, and Colombia (specifically Bogota), highlight how design aimed at changing the behavior of those that need create the enabling environment is critical for change too. All of the above can be found in Part II of this report.

Rather than relying on a single behavioral insight, draw on multiple psychological and social principles to address actors' various barriers and motivations.

Applying behavioral insights to a given behavior change challenge involves connecting the findings from the behavioral context analysis to principles documented in the behavioral and social sciences as being important enablers and drivers of behavior. This can be contrasted with the common practice of relying on the designer's intuition for what might drive behavior change in a particular context. This also does not mean applying the same principles in every context. Rather, it means identifying which principle is applicable given the behavioral context analysis. For more information on engaging in this process, please refer to Part I of this report, which includes information on how various frameworks approach mapping behavioral and social science principles to what will change behavior. It is important to note, however, that the principles in those frameworks far more frequently drawn from the behavioral rather than social sciences. This is not to suggest that social science principles are not applicable, but behavioral science attempts to identify domain-general principles, whereas the social sciences tend to focus on how different contexts make different intervention features more or less relevant. Within a given context, however, the social sciences often provide specific applicable principles. For example, if a designer is facing a common pool resource problem, Ostrom's eight Design Principles for Common Pool Resource Institutions would be particularly relevant. Once a program designer has identified their area of intervention, they should then identify which social science principles correspond to that specific domain.

At this stage, many programs make the mistake of focusing on a single behavioral or socio-ecological insight. This myopic focus provides a single point of program failure, as well as fails to recognize the differing social positions actors, their individual differences in motivation, and the fact that even a single actor often faces multiple distinct barriers and motivations which making a behavioral choice.

There can be different barriers and motivations within a given actor-behavior pair (e.g., farmers using compost). For example, the social structure could mean that male farmers have greater authority in managing waste for composting. This insight might lead an intervention designer to conduct additional programming to change the social dynamic of women's authority in this context to afford them the ability to choose whether to compost in their field. At an individual level, behavioral context analysis may also show that aversion to ambiguity is a major barrier

to adopting this novel practice. A program that only addresses the social-structural insight of gender differences or only addresses the psychological insight of aversion to ambiguity is likely to be insufficient to tackle this challenge. Instead, effective programs knit together a web of behavioral and socio-ecological insights to motivate each element of the intervention.

The context will determine what behavioral and social science insights to incorporate. The case studies in Indonesia and Mexico compellingly integrate a suite of psychological and socio-behavioral insights to develop improved behavior change programming.

Throughout the intervention development process, encourage prototyping interventions, and test those prototypes with the target actors prior to scaling implementation.

While developing an intervention based on an empirically supported psycho-social theory of change is far more likely to render an effective behavior change intervention, it is impossible to perfectly forecast exactly how actors will engage with each intervention component. This means it is critical to test those intervention components with the target actors in order to reduce the possibility of deploying an intervention at scale that then later proves to be ineffective. There are various scales at which this testing can occur: from presenting prototypes (small scale versions, generally at a far lower cost and effort) of intervention components and getting direct feedback in something like a focus group, all the way to fully piloting interventions in target communities before finalizing them for deployment at scale. Where on this spectrum testing falls should depend on the confidence designers have in the effectiveness of a particular component in the target context and the costs of an intervention not being successful by not pursuing more extensive testing. Testing and revising imply that intervention development is necessarily non-linear. Instead, testing and intervention development represent an iterative process—each feeding into the other until the intervention is ready to be launched at scale.

Testing can sometimes be perceived as a costly or risky step in the intervention development process, but it is far more often net cost saving. While testing adds direct costs, it allows a designer to avoid the far more costly error of implementing an unsuccessful intervention. Similarly, while a test may fail in the sense that it shows that a particular component is unlikely to achieve its desired result, this type of failure massively reduces the risk of a large-scale failure if that component was taken to scale.

For more information on testing, designers may wish to refer to the World Bank and Rare frame-works in Part I of this report. The case studies in Kenya and Brazil also present clear demonstrations of how testing intervention prototypes can dramatically improve program effectiveness. The water crisis case study in Colombia, however, gives a cautionary perspective on how a lack of adequate testing may result in significant costs or delayed action.

Criteria for Monitoring and Evaluation

Building a monitoring and evaluation framework for behavior change interventions that can generate actionable information for live program decisions, accurate evaluation, adaptive management, and understanding program generalizability.

Monitor and evaluate the intervention not just against environmental outcomes, but also behavioral, psychological, and social outputs

Selecting indicators to monitor should be derived from a program's ToC. Incorporating psycho-logical and social states into the PS-ToC allows for these indicators to be incorporated into pro-gram monitoring and evaluation. Monitoring these psychological and social states allows for a program designer to assess the effectiveness of the various component mechanisms for change as well as behavioral measures and intended environmental outcomes. This allows designers to take a more nuanced approach to adaptive management, and to improve on future program development.

While assessing psychological and social indicators can be important for assessing the effectiveness of intervention elements in achieving intermediate outputs, they can also be important program outcomes in and of themselves. Many programs assume that achieving environmental outcomes will naturally lead to positive social outcomes for affected communities. However, there are numerous documented cases where positive social outcomes failed to materialize, or worse, where interventions resulted in harm to the communities that rely on the environmental resources that were targeted. In cases where interventions may affect the social welfare of a community—occurring in the large majority of community-based behavior change programs—social outcomes need to be evaluated with the same attention and rigor as environmental outcomes.

Collecting monitoring data on psycho-social intermediate states and outcome indicators provides unique data for decision making in monitoring, evaluation, and generalization. Monitoring data allows designers to more accurately assess the underlying mechanism for why a program works, which can be incorporated into adaptive management decisions over the lifecycle of a program. When collected at a high-enough frequency, these data can also be used for dynamic program-ming on a site-specific level—meaning that data feeds back into a program while it is being delivered to inform live decision making on improvements to delivery, reductions in program cost, and increases in effectiveness. Additionally, as psycho-social indicators provide insight into why a program functions, they can be used to inform the generalizability of an intervention to novel contexts, making the scaling out of a program far more likely to succeed. To find more details on these applications of psycho-social indicators, please refer to the Framework for Indicator Monitoring and Adaptive Management section of Part III of this report.

A program's PS-ToC will determine the specific behavioral, social, and psychological indicators relevant for monitoring and evaluation. However, for additional general guidance on the selection and measurement of behavioral, social, and psychological indicators, please refer to Part III of this report.

Evaluate the program against a valid counterfactual to determine the causal effect of the intervention.

Understanding whether a program achieved its desired outcomes, be those be environmental, behavioral, psychological, or social, allows future programs to learn, and effective programs to be scaled out. However, frequently program outcomes are assessed in a way that it is hard to justify any causal claim about the effect of the intervention.

Causal reasoning depends on comparing against a counterfactual, which asks, what would have happened if no intervention had taken place? Without some understanding of what the counter-factual is, there is no way to determine if a change in the value of an indicator was due to an intervention, or anything else in the behavioral or social context that might have impacted it. Randomization is a commonly-used technique for creating a counterfactual; some units, such as individuals or communities, are randomly assigned to receive an intervention, and others are not. The difference between the outcomes of these two groups can thus reasonably be understood as causal—that is, designers can be confident that it is the intervention that caused any changes they observed.

When randomization is not possible, there are various alternative techniques. These include matching and difference-in-difference, which allow for the estimation of an intervention's causal effect, though they require additional assumptions. Critically, the use of simple, pre-post comparison is nearly never sufficient for implementers to reliably estimate the effect of an intervention because many factors other than the intervention itself may have influenced the intervention's outcome indicators.

For more details on the options available for evaluation designs, including those identified here, please refer to Frameworks for Evaluating Changes in this report's Part III, Indicators section.

Continue evaluation beyond the program's end to ensure that environmental out-comes, as well as the psychological and social outputs, are sufficiently durable.

Particularly in the environmental sector, the impacts we aim to achieve often last far beyond the termination of program activities. However, for many programs, the final evaluation of that impact occurs at the time the program concludes. While this is problematic for any environmental program, it is a particular issue for behavior change programs that rely on sustained adherence to the target behavior to achieve maximum impact.

This misalignment between a program's impact and its evaluation timeline creates a perverse incentive for program designers. Programs that achieve quick but unsustainable changes in behavior are seen as relative successes, whereas those that rely on scaling deep are often seen less favorably than they deserve. For example, social norm change programming can create self-reinforcing norms, but those take significant time to achieve.

Programs should, therefore, match their evaluation timeline with the timeline of the expected impact of the program, rather than the duration of the program itself. This is often most effectively done through the use of proxy measures for key indicators that can rely on either remote or administrative data collection, making continued impact evaluation more cost-effective. Such evaluation allows for more effective data for decision-making by allowing for the entire impact of a program to be incorporated into decisions to scale the program out.

PART I:

Review of Behavior Change Approaches

Introduction

There are many approaches, tools, and perspectives on how to change behavior. Behavior change approaches are particularly valuable because they help to distill the many insights about human behavior into something concrete and accessible as well as offer a series of steps for designing behavioral solutions. In other words, it is not enough to understand behavioral and social science principles; one must also know how to put those into practice (Wendel, 2020).

In the last decade, the field of ‘behavioral design’ has emerged as a blend of insights from the design and behavioral and social science fields (Datta & Mullainathan, 2014; Wendel, 2020). The field of design provides a useful sequencing strategy for generating solutions and an explicit focus on centering the target actor. It further provides a creative and innovative mindset for designing solutions that go beyond individuals and that seek to transform larger systems (Reid & Schmidt, 2018). Design thinking and Human-Centered Design are especially popular among people who do this work. The behavioral and social sciences layer on research about people’s motivations and barriers for behavior change and intervention strategies that could prove most impactful. For many designers today, behavioral design also involves an explicit focus on ‘behavior change for good,’ meaning solutions that promote behaviors that support human wellbeing (Salzer, 2020).

In this section, we describe the similarities and differences between six leading behavior change approaches. We also offer best practice recommendations for selecting a formal behavior change approach. Then we provide a more in-depth review of each of the six behavior change approaches, focusing on the insights, methodologies, and resources for creating effective behavioral interventions. The selection criteria for the six approaches in this section were the following (See Appendix A for the complete list of approaches we considered):

- Behavioral: The approaches all have a specific framework for understanding and analyzing insights from the behavioral and social sciences
- Design-focused: The approaches detail a methodology for designing behavior change interventions
- Environmental: The approaches were relevant for environmental work or had been applied to behavior change problems in any of the GEF’s focus areas

There is significant overlap among these approaches in terms of their understanding of the behavioral and social sciences and in their design methodologies, yet each has unique strengths and insights. For each approach, we highlight its origin, the behavioral principles or framework used to design solutions, the design steps or processes, as well as any known applications of the approach in environmental contexts, particularly in GEF focus areas. We then evaluate each approach’s strengths and weaknesses based on the following criteria:

- Variety and strength of behavioral principles
- Comprehensiveness of design steps (from understanding the problem to assessing and monitoring outcomes)
- Consideration of behavior change durability and evaluation
- Ability to work among diverse sets of actors and at different scales of governance
- Integration of the socio-ecological context for a behavior change problem
- Emphasis for design steps to be non-linear, adaptive, and iterative
- Consideration of ethics and power in designing for behavior change

We have used a dot system (● ● ○) to label how strong an approach is for each criterion: dark blue indicates the criterion is met comprehensively and missing no elements; light blue indicates the approach is not meeting one core element of the criterion, and white indicates that the approach is not meeting multiple elements of the criterion. These evaluations can be used to assess the relative emphasis each approach gives to elements of designing for behavior change and identify where there is room for ongoing improvement in the field.

Here is a summary table of how each of the approaches fared on the seven criteria:

| Criteria | Behavior Change Approach | | | | | |
|--------------------------------------|--------------------------|-------|------|------|------------|---------|
| | BIT | COM-B | OECD | Rare | World Bank | Ideas42 |
| Strength of behavioral principles | ● | ● | ● | ● | ● | ● |
| Comprehensive design steps | ● | ● | ● | ● | ● | ● |
| Durability and evaluation | ● | ● | ● | ● | ○ | ● |
| Diverse actors and scales | ● | ● | ● | ● | ● | ● |
| Socio-ecological context integration | ● | ● | ● | ● | ● | ● |
| Emphasis on non-linearity, iteration | ● | ● | ● | ● | ● | ● |
| Ethics and power consideration | ○ | ● | ● | ● | ● | ● |

Observations and Recommendations For Selecting a Behavior Change Approach

Based on the criteria we used to evaluate behavior change approaches, we observed several patterns in terms of their similarities and differences. We also present a number of recommendations for what comprises a strong behavior change approach for each of the above evaluation criteria.

Behavior Change Approach Similarities and Differences

The six chosen approaches have more similarities than differences. They have all been highly influenced by the latest behavioral and social science as well as design thinking. Additionally, all of them are strong in their applicability to diverse sets of actors and scales of governance. Beyond that, there are minor differences in the behavioral principles and design steps they use to guide the design of behavior change interventions.

Behavioral principles

Some approaches describe their behavioral principles and strategies as broad categories (e.g., automatic thinking, social influences, choice environment); others are more specific and target behavioral or social principles directly (e.g., messenger effect, salience, loss aversion). Some approaches distinguish between rational and irrational forms of thinking in the way they describe behavioral insights, and some also distinguish between ‘fast’ and ‘slow’ or System I and System II thinking. These are categorical distinctions that could create false dichotomies with regards to the way in which decisions are made. These types of thinking may lead to rigid boundaries around the selection of certain strategies (Osman, 2004). For example, the use of emotional framings can appeal to our ‘automatic’ processing system given how we respond to emotional events, but it can also be intensely rational in pursuing individual needs (e.g., seeking safety from a threat; Lerner et al., 2015). Finally, most approaches include behavioral principles that represent ‘new’ insights from behavioral science that depart from the more traditional rational actor models. This also means that they tend to exclude strategies such as providing material incentives, information, or enacting regulations. In contrast, the COM-B and Rare approaches include these strategies to highlight their value and role in changing behavior, even if not on their own.

Design steps

The design steps for all approaches are very similar to one another. They all include a step to define the target behavior and target outcomes, understand and analyze the motivations and barriers of the target actors, design solutions, and implement those ideas. They differ where the steps ‘start’ and ‘end’; some approaches are more explicit about spending time early on to gather details about the context in which a behavior occurs. Others spend more time detailing how to test, evaluate, and monitor interventions. Approaches also vary in where they incorporate iteration as part of the design process and how mindful they are of power and ethics throughout an intervention’s design.

Behavior Change Approach Best Practices

No single approach is perfect across all criteria; instead, each presents unique strengths and weaknesses that we can utilize and learn from. When choosing a behavior change approach, whether one of the six we selected or an amalgamation of several, we recommend designers aim to maximize each of these criteria by looking for the following best practices:

Variety and strength of behavioral principles

Behavioral principles serve as the behavioral and social science framework through which designers understand and analyze the motivations and barriers that their target actors have relative to behavior change. They also help to design effective interventions. There is no 'right' framework for this, but some approaches have frameworks that are more holistic, evidence-based, and better able to categorize and understand the insights designers gather about the target actor. Look for behavioral principles that can account for internal and external, personal and social, and effortful and unconscious reasons for change. Some behavior change approaches make no explicit mention of strategies like providing information, material incentives, or enacting rules. Instead, these approaches either ignore those strategies altogether or see them as needing to be paired with something like social norms messaging or choice architecture. The risk here is that effecting behavior change can be as simple as providing an incentive, but it can also require a complex combination of different strategies. Ultimately, what is most important is that the strategy designers choose is informed by the data they have collected about their target actors, the target behavior, and actors' position in the relevant socio-ecological context.

Comprehensiveness of design steps

The design steps of a behavior change approach provide implementers and designers with a logical sequence of steps to design a behavior change solution. There is no 'right' set of steps. Even so, we find that most approaches share in a number of core steps that ensure designers are effective. Firstly, designers need to develop an understanding of the socio-ecological context in which current and target behaviors take place. This also means identifying systems, institutions, structural barriers, and enabling factors that may affect the intervention's capacity for behavior change. Secondly, designers need to identify the stakeholders, target actors, and target behavior(s) related to the behavioral outcome they want to achieve. This can also be revised after the third step, which involves qualitative and/or quantitative research on the target behavior and target actor, usually via the tools of social sciences (e.g., ethnographic analysis, surveys, participant observation, etc.). After collecting data, the fourth step involves analyzing data with a behavioral lens. Specifically, designers need to identify the key motivations and barriers driving the actors' behaviors and the behavioral strategies most likely to encourage change. This also requires acknowledging the socio-ecological system from the first step and incorporating any relevant insights that could affect an actor's ability to change. In this step, designers may also develop a psycho-social theory of change (PS-ToC) to guide their intervention.

Design steps now shift from collecting and analyzing data to prototyping and testing an intervention. Testing and revising the intervention multiple times provides valuable insights about whether the intervention is working and what designers could do to improve it before taking it to scale. This may involve another round of quantitative and/or qualitative research, letting designers fill the gaps in their understanding of what will change behavior. Once designers are confident that their intervention is effective in its test phase, they can move to implement it at scale and assess and monitor the results. This involves having a rigorous evaluation strategy that assesses the true impact of the intervention relative to a counterfactual: in other words, a representation of what would have happened if no intervention had taken place. It is also important that designers' design journey not stop there; they should continue to monitor how the intervention may need to adapt and how the target behavior changes over time. An overview of all design steps appears in the figure below.

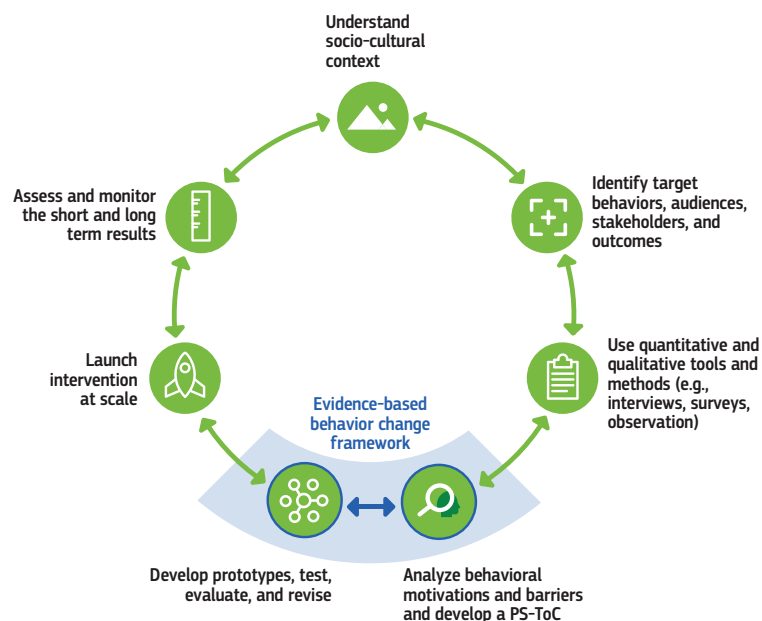


Figure 1: An overview of design steps for behavior change interventions

Consideration of behavior change durability and evaluation

Many behavior change interventions are designed to work for a specific time. We cannot assume that these changes will be durable during or beyond the intervention period when there is no formal monitoring and evaluation strategy. It is important to identify that the intervention is the reason behind a change in behavior and to know that this change will persist. There is, however, a systematic and consistent gap in behavior change approaches relative to how they assess the ongoing impacts of interventions. Additionally, some behavioral design approaches focus heavily on the intervention design, yet do not spend adequate resources on developing an evaluation framework. In cases where rigorous assessment is feasible, it is important to invest time and energy in this framework as a way to measure the effectiveness of the solution and its ability to scale. This means that designers should identify their key indicators and outcomes, determine how they will measure those indicators, and then decide how they will evaluate changes to those indicators. See Part III: Behavior Change Indicators for further recommendations and details for building a robust evaluation framework into the design process.

Ability to work among diverse actors and scales of governance

A general behavior change approach should be applicable to any behavior change problem, regardless of the actors or scale involved. Yes, some actors may be harder to access for various reasons (e.g., government officials, remote communities), but that does not mean that they should not be considered in intervention design. Moreover, effective behavior change interventions tend to involve actors beyond the direct actor. A behavior change approach should be able to plan for and accommodate the roles that other stakeholders may play as part of a given solution.

Working on behavior change at larger scales can be tricky; there are usually more stakeholders involved and contextual factors to consider. Yet, designers should also be able to reapply the same steps they used above to extend their intervention past a narrow set of actors. Addressing systems-level change with a behavioral lens involves working on the system piece by piece, where each is its own behavioral problem.

Integration of understanding the socio-ecological context

An understanding of the socio-ecological context in which a behavior occurs is essential information for the design of any behavior change intervention. A behavior change approach should dedicate adequate time to mapping out

the systems surrounding a behavior. It is likely that these systems will have a significant impact on the intervention options available, whether behavior change is feasible, and if so, for whom. Questions such as ‘Which behaviors are related to my environmental outcomes?’, ‘What is the network of actors connected to these behaviors?’, and ‘Where do I have the most impact and ability to make change?’ are important starting points. Goddard et al.’s (2016) values, rules, and knowledge framework is one way to capture and understand the systemic barriers, institutions, and principles that create the context for behavior change. In understanding this broader context, designers learn about the options that are available to different actors, about who carries decision-making power, and about the behaviors that could reinforce or challenge existing power structures.

Oftentimes behavioral designers are not embedded or a part of the context in which they wish to make change; as a result, they bring a number of information gaps and assumptions into their work. Self-aware designers benefit from noting upfront any personal assumptions before they take the time to understand the behavior and context. Another solution is for designers to co-design with the stakeholders of the behavioral problem to ensure they have an informed perspective at each step of the process. With or without these stakeholders, designers should strive to gain a deep understanding of the behavioral problem at hand before generating any solutions.

Emphasis for design steps to be non-linear, adaptive, and iterative

A good behavior change approach should have the flexibility to be iterative and repeat steps as needed. It is extremely unlikely that designers will get everything ‘right’ with their first intervention idea. Building in the expectation and time to learn more, make revisions, and complete multiple rounds of testing increases the likelihood that designers will meet their behavior change goals. A stepwise process is important, but designers should not feel as if it is restrictive and linear. They should be able to stop, go back, and repeat steps throughout the process. They may even choose to return to the first step after implementing and assessing their intervention in a series of full- or half-cycles, as exemplified by the World Bank’s approach.

Consideration of ethics and power in designing for behavior change

Any behavior change approach should strive to maintain individual rights and help people to pursue their own behavior change goals. In the design steps, a behavior change approach can encourage designers to reflect on every decision they make: for example, what it means to focus on specific behaviors and actors in a given context, who benefits, and who could be negatively impacted in intended or unintended ways. It also means acknowledging and addressing the biases that exist within designers as well as those reinforced by the dynamics of actors in the behavioral system. Environmental designers often come with their own (often academic) backgrounds and perspectives that are different from the local expertise of their target actors. The more time designers take to learn about their target behaviors through the lens of their target actors, or co-design solutions with local stakeholders, the more opportunity there is to shift the power imbalances in intervention design. The OECD’s approach offers a checklist at each step of the design process to help assess the ethical considerations of an intervention. The World Bank also conducted research on its own staff to better understand the biases of development designers.

Any behavior change approach should further be mindful of ethics and power in its use of behavioral principles. Research on behavioral economics, choice architecture, and ‘nudging’ has gotten a lot of attention and criticism from scientists and the public. These tools can be powerful and have indeed demonstrated major impacts on a wide range of topics. Nevertheless, it is important to recognize that these are just one part of a behavior change toolkit. It is very unlikely that a singular strategy would serve as the ‘silver bullet’ to any one behavior change problem. Unfortunately, some nudging or choice architecture interventions have been described as universally applicable strategies—a perspective that can have disempowering results for many target actors. No matter the strategy, behavior change interventions that use forms of coercion and manipulation create an innate power imbalance and should be avoided. A behavior change approach that aims for real, fair, and durable environmental change should leverage the target actor’s own insights in intervention design.

Review of Selected Behavior Change Approaches

In this section, we describe six different behavior change approaches applicable to the environmental field and analyze their strengths and weaknesses relative the evaluation criteria mentioned above.

The Behavioral Insights Team: MINDSPACE, EAST, TESTS

Background

The Behavioral Insights team created MINDSPACE in 2010—simplifying it to EAST in 2012 (Dolan et al., 2010; The Behavioural Insights Team, 2012). These approaches were initially designed to guide policy-makers and government officials into leveraging behavioral theory. They have since expanded and now also cater to NGOs and private firms.

Behavioral principles

MINDSPACE and EAST are acronyms for behavioral principles and strategies used in the design of behavioral interventions:

MINDSPACE

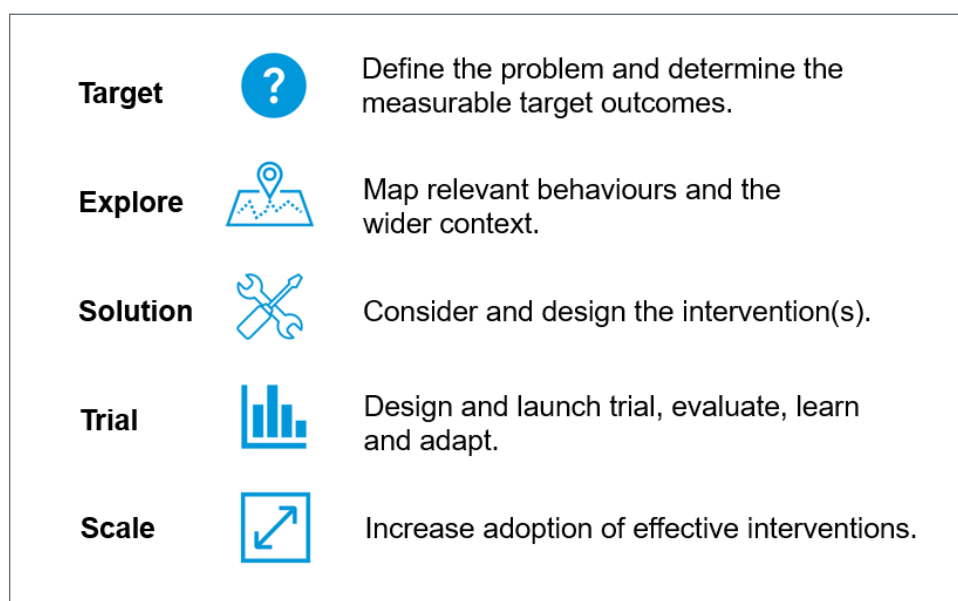
- Messenger: We are heavily influenced by who communicates information
- Incentives: Our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses
- Norms: We are strongly influenced by what others do
- Defaults: We “go with the flow” of pre-set options
- Salience: Our attention is drawn to what is novel and seems relevant to us
- Priming: Our acts are often influenced by sub-conscious cues
- Affect: Our emotional associations can powerfully shape our actions
- Commitments: We seek to be consistent with our public promises, and reciprocate acts
- Ego: We act in ways that make us feel better about ourselves

EAST

- Easy: Harness the power of defaults; Reduce the ‘hassle factor’ of taking up a service; Simplify messages
- Attractive: Attract attention, Design rewards and sanctions for maximum effect
- Social: Show that most people perform the target behavior; Use the power of networks; Encourage people to make a commitment to others
- Timely: Prompt people when they are likely to be most receptive; Consider the immediate costs and benefits; Help people plan their response to events

Design steps

The Behavioral Insights Team uses a methodology with the acronym TESTS, and the steps are defined in the graphic below (Green et al., 2019):



Source: (Green et al., 2019)

Environmental applications in GEF focal areas

- Reviews by Palm-Forster et al. (2019) and Shreedhar (2020) summarize behavioral interventions applied to agri-environmental and conservation contexts (developed in the academic literature) according to the MINDSPACE framework.
- **Wildlife Conservation:** The aim of the project was to encourage volunteers to take a more active role in the local government's 'Eco Village' program (The Behavioral Insights Team, 2019).
- **Chemicals and Waste Management:** The research introduced two different initiatives to encourage the use of reusable lunch containers and reduce single-use plastic (The Behavioral Insights Team, 2020).
- **Climate Change Mitigation and Adaptation:** This study involved distributing a new thermostat that makes energy saving attractive by nudging consumers to "collect green leaves" (The Behavioral Insights Team, 2017).

Strengths & weaknesses

- **Variety and strength of behavioral principles:** Both MINDSPACE and EAST simplify a variety of behavioral principles into memorable mnemonic devices. The frameworks do not include more standard or traditional strategies such as enacting laws or providing information.
- **Comprehensiveness of design steps:** The TESTS approach has a comprehensive set of steps, ranging from defining the problem and its possible outcomes, to evaluating and scaling the possible solution(s).
- **Consideration of behavior change durability and evaluation:** MINDSPACE and EAST both describe how certain strategies are more durable than others, such as providing reminders (less durable) vs. establishing social norms (more durable). TESTS also includes explicit guidance for testing and evaluating interventions yet offers less on monitoring outcomes over time.
- **Ability to work among diverse actors and scales of governance:** The Behavioral Insights team was founded

with the aim of incorporating behavioral insights into bureaucratic processes, public service, communications, and policy. Since its founding, their scope of work has broadened to include NGOs and private firms. Their projects show good representation of actors and outcomes at the household, organization, and government level.

- **Incorporates socio-ecological contextual analysis:** TESTS emphasizes the importance of understanding the context in which a behavior takes place as well as the larger systems that shape it. Still, it does not provide tools or guidance for how to analyze these systems, or about how to establish checks so that such an analysis appears in solution design.
- **Emphasis for design steps to be non-linear, adaptive, and iterative:** TESTS explicitly mentions how the steps of the process are non-linear and involve iteration and feedback loops.
- **Consideration of ethics and power:** This framework considers the role of designers as well as how different actors may have the agency to change their behavior. It does not provide concrete recommendations for addressing the bias of intervention designers, nor the many ways that power asymmetries may influence the various steps of the design process.

For more information on the Behavioral Insights Team's approach, explore the following resources:

Dolan, P., Hallsworth, M., Halpern, D., King, D., & Vlaev, I. (2010). *MINDSPACE: Influencing behaviour for public policy* [Monograph]. Institute of Government. <http://www.instituteforgovernment.org.uk/publications/>

Green, K., Williamson, K., Park, T., & Reiner, C. (2019). Behavior Change for Nature: A Behavioral Science Toolkit for Practitioners. Rare. <https://behavior.rare.org/wp-content/uploads/2019/10/2019-Behavior-Change-for-Nature-Report-digital.pdf>

The Behavioural Insights Team. (2012). EAST: Four simple ways to apply behavioural insights. https://www.bi.team/wp-content/uploads/2015/07/BIT-Publication-EAST_FA_WEB.pdf

COM-B and The Behavior Change Wheel

Background

Michie et al. (2011) developed a set of guidelines for designing behavior change interventions based on a review of 19 frameworks and 7 policy categories. The COM-B model describes a person's physical and psychological capability (C), physical and social opportunity (O), and automatic and reflective motivation (M) as the core causes of behavior (B) at any given moment. The Behavior Change Wheel presents the COM-B model as the center circle and concentric circles of intervention functions (or behavioral principles/strategies) and policy categories.

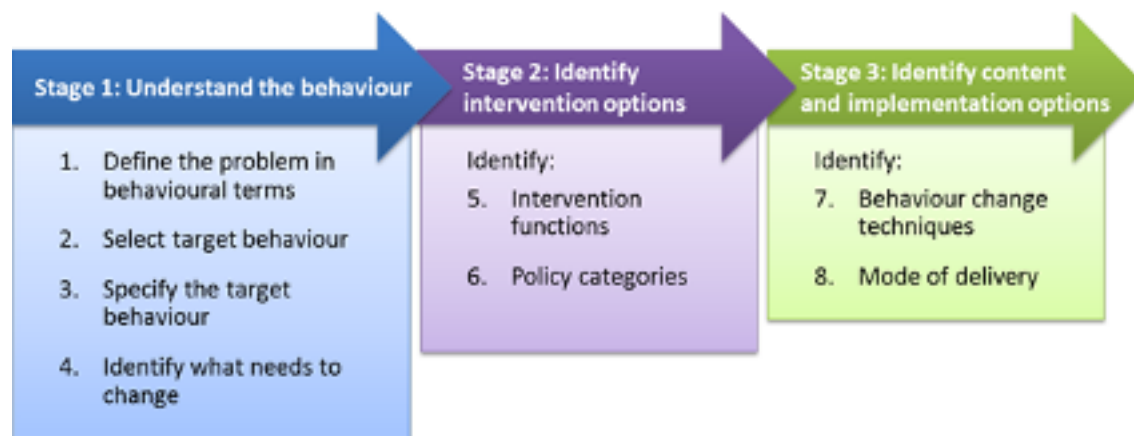
Behavioral principles

The intervention functions in COM-B and the behavior change wheel serve as the behavioral principles or strategies employed in this approach:

- Persuasion: Using communication to induce positive or negative feelings, or to stimulate action
- Education: Increasing knowledge or understanding
- Restrictions: Using rules to reduce the opportunity to engage in the target behavior (or to increase the likelihood of engaging in the target behavior by reducing the opportunity to engage in competing ones)
- Environmental restructuring: Changing the physical or social context
- Modeling: Providing an example for people to aspire to or imitate
- Enablement: Increasing means/reducing barriers to increase capability or opportunity
- Training: Imparting skills
- Coercion: Creating an expectation of punishment or cost
- Incentivization: Creating an expectation of reward

Design steps

COM-B and the Behavior Change Wheel approach use the following steps over three stages for designing interventions (Michie et al., 2014):



Source: Michie, S., Atkins, L., & West, R. (2014)

Environmental applications in GEF focal areas:

- **Wildlife Conservation:** The COM-B framework was used in a review to analyze the potential routes to encourage higher welfare food choices to increase the well-being of farm animals (Cornish et al., 2019).
- **Chemicals and Waste Management:** The COM-B framework guided qualitative research that aimed to develop behavioral interventions to encourage recycling behaviors (Gainforth et al., 2016). These interventions are yet to be tested.
- **Climate Change Mitigation and Adaptation:** The framework was used to guide a game design that aimed to encourage energy conservation (Wells et al., 2016).
- **Land Degradation:** The COM-B framework was applied to reveal factors that facilitate and impede the adoption of sustainable farming in the face of the threat of invasive species in Austria (Kropf et al., 2020).

Strengths & Weaknesses

- **Variety and strength of behavioral principles:** This approach covers the full range of strategies, including more traditional ones like rules and regulations, information, incentives, in addition to emotional, social, and environmental types of motivations.
- **Comprehensiveness of design steps:** The design steps focus heavily on understanding context and designing appropriate solutions. While the delivery of interventions is also emphasized, there is less explicit detail on testing, launching, or evaluating interventions.
- **Consideration of behavior change durability and evaluation:** This approach describes the importance of rigorous evaluation and measuring change over time but does not include explicit guidance about how to do these steps.
- **Ability to work among diverse actors and scales of governance:** This approach is explicit that it was designed to work across local, regional, and national scales as well as with different sets of actors (ranging from single individuals to entire communities). It is highly sensitive to the scale of intervention needed, especially within the way it defines different policy categories.
- **Incorporates socio-ecological contextual analysis:** This approach is strong in its contextual analysis and in the way it sees any behavior as part of a greater system of actors and structures. Still, it lacks an explicit focus on socio-ecological barriers, focusing more on physical or cognitive barriers to change.
- **Emphasis for design steps to be non-linear, adaptive, and iterative:** This approach is flexible for making revisions to one's design but is not explicitly non-linear or iterative.
- **Consideration of ethics and power:** This approach takes great care in understanding the social context (e.g., actors and system) in which a behavior operates and includes criteria for assessing the acceptability, side effects, safety, and equity of an intervention.

For more information on the Behavior Change Wheel approach, explore the following resources:

Michie, S., Atkins, L., & West, R. (2014). *The Behaviour Change Wheel: A Guide to Designing Interventions*. Silverback Publishing.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*: IS, 6, 42. <https://doi.org/10.1186/1748-5908-6-42>

OECD: BASIC

Background

The OECD created the BASIC toolkit in 2019 for policymakers that includes tools, methods, and guidelines for designing behavioral interventions at all stages in a policy cycle (OECD, 2019).

Behavioral principles

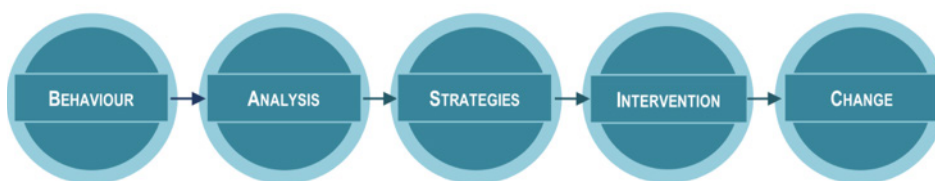
This approach uses the following 'ABCD' principles to guide interventions:

- Attention: People's attention is limited and easily distracted
- Belief formation: People rely on mental shortcuts and often over/underestimate outcomes and probabilities
- Choice: People are influenced by the framing of choices as well as their social and situational context
- Determination: People's willpower is limited and subject to psychological biases

Design steps

The acronym, BASIC, is a way to remember the steps in the OECD's behavioral design process:

1. Behavior: Identify and target the crucial behavioral aspects of the policy problem.
2. Analysis: Scrutinize target behaviors through the lens of behavioral science.
3. Strategy: Identify and conceptualize behaviorally informed policy solutions.
4. Intervention: Design experiments to evaluate the effectiveness of the strategies.
5. Change: Plan for implementation, scale, monitoring, evaluation, maintenance, and dissemination of results.



Source: OECD (2019)

Environmental applications in GEF focal areas

Not yet available due to this approach having been developed only in the last year.

Strengths & weaknesses

- **Variety and strength of behavioral principles:** ABCD is a helpful acronym, yet it is a limited framework for categorizing behavioral motivations and barriers. It both does not cover strategies such as incentives and information, and some of the principles appear to be overlapping.
- **Comprehensiveness of design steps:** This approach explicitly covers all of the important steps in the design process, from identifying behaviors and outcomes, to implementation, to the evaluation of results. Still, the step on collecting data on motivations and barriers lacks detail. This approach jumps from identifying a behavior and its context, to analyzing motivations and barriers.

- **Consideration of behavior change durability and evaluation:** The toolkit specifically mentions the importance of measuring, evaluating, and sharing the impacts of an intervention.
- **Ability to work among diverse actors and scales of governance:** Designed for the policy sector, this approach is flexible and functions at scales that range from households to governments.
- **Incorporates socio-ecological contextual analysis:** The authors emphasize the need to understand the context of a behavior as well as engaging those stakeholders that are related to it. This is done to better understand the system in which the behavior operates. The authors also offer system mapping tools, but those focus mainly on actors and behaviors, not on socio-ecological factors. Due to this approach's emphasis on ethics, designers are further prompted to consider feasibility and the structural barriers that could impede their options.
- **Emphasis for design steps to be non-linear, adaptive, and iterative:** The process is iterative and offers the ability to refine outcomes and solutions while gathering more information.
- **Consideration of ethics and power:** This framework uniquely prioritizes ethical considerations at each phase of the design process.

For more information on the OECD approach, explore the following resource:

OECD. (2017). *Behavioural Insights and Public Policy: Lessons from Around the World*. OECD Publishing. https://www.oecd-ilibrary.org/governance/behavioural-insights-and-public-policy_9789264270480-en

Rare: Behavior-Centered Design and Levers of Behavior Change

Background

Rare's Center for Behavior & the Environment developed both a behavior change framework and an approach for environmental designers based on lessons learned during their many years of behavior change work in conservation (Green et al., 2019; Rare, 2020).

Behavioral Principles

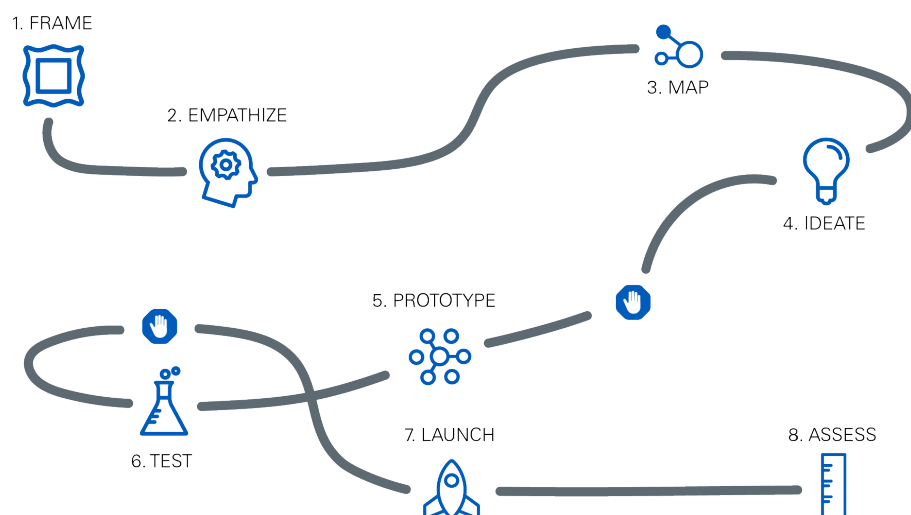
Rare uses six levers of behavior change that can be pulled individually or in combination to achieve different effects when designing interventions:

- Emotional Appeals: Using emotional messages to drive behavior
- Social Influences: Leveraging the behaviors, beliefs, and expectations of others
- Choice Architecture: Changing the context in which choices are made
- Rules and Regulations: Enacting rules that promote or restrict a behavior
- Material Incentives: Increasing or decreasing real or perceived costs, time, or effort for doing a behavior
- Information: Providing information about what the target behavior is, why it matters, and how to do it

Design steps

Rare's design approach is called Behavior-Centered Design and uses the following steps:

- Frame: Frame the conservation challenge to understand the target behavior, target actor, and describe their context.
- Empathize: Gain deep insights into the target actor's relationship with the target behavior, including the motivations and challenges they face.
- Map: Organize insights about your target actor into behavioral motivations and challenges.
- Ideate: Generate, group, and prioritize intervention ideas.
- Prototype: Select your best intervention idea and develop a prototype (small-scale version) that captures its essential features.
- Test: Test your prototype with your target actor and make revisions based on their feedback and your own experience.
- Launch: Plan and launch your full-scale intervention and think about evaluation and impact metrics.
- Assess: Assess the impact of your intervention and reflect on potential improvements.



Source: Green et al. (2019)

Environmental applications in GEF focal areas:

- **Biodiversity conservation:** Fishers were given a tool to measure the size of lobsters easily (to prevent them from fishing small lobsters) and a communication campaign 'The Size Matters' to educate fishers about sustainable fishing methods (Green et al., 2019).
- **Biodiversity conservation:** Rare developed a Theory of Cooperative Behavior Adoption that provides guidance for addressing common-pool resource dilemmas in the environmental field, including water and land management, and draws upon the BCD approach and behavior levers framework (Thulin, 2020). Rare's Fish Forever program also uses this theory to design community-based interventions.
- **Land management:** Rare launched a social marketing campaign in Mexico to join a Payments for Ecosystem Services program and conserve forest habitat (Green et al., 2013)
- **Water management:** Rare's watershed program in the Andes region facilitated communication and cooperation between upstream and downstream users to ensure clean water for everyone (Rodríguez-Dowdell et al., 2014).

Strengths & weaknesses

- **Variety and strength of behavioral principles:** The levers of behavior change cover common and traditional strategies, like information and incentives, as well as the latest findings in behavioral insights for emotional, social, and choice architecture-based motivations.
- **Comprehensiveness of design steps:** The approach considers a full range of steps from identifying the problem to evaluating interventions in its intervention design process
- **Consideration of behavior change durability and evaluation:** This approach helps designers evaluate their interventions but offers less guidance on assessing durability or on the monitoring of interventions over time.
- **Ability to work among diverse actors and scales of governance:** This approach has mostly been used to design interventions for households and communities, yet it can easily be applied to more diverse sets of actors and scales (whether local, regional, or national).
- **Incorporates socio-ecological contextual analysis:** This approach encourages designers to analyze the broader context of behavioral outcomes (particularly as they relate to sustainable behavior change) yet does not have specific tools for building socio-ecological variables into solution design.
- **Emphasis for design steps to be non-linear, adaptive, and iterative:** This approach clearly emphasizes the importance of revising solutions over many iterations.

- **Consideration of ethics and power:** This approach considers the kinds of biases that designers bring to their work, the broader systems in which behaviors occur, as well as how to best involve target actors meaningfully during solution design. Yet, it is less clear how these considerations fit into every stage of the design process.

For more information on Rare's approach, explore the following resources:

Green, K., Williamson, K., Park, T., & Reiner, C. (2019). *Behavior Change for Nature: A Behavioral Science Toolkit for Practitioners*. Rare. <https://behavior.rare.org/wp-content/uploads/2019/10/2019-Behavior-Change-for-Nature-Report-digital.pdf>

Rare. (2020). Behavior Change for the Environment – Rare. <https://behavior.rare.org>

World Bank: Mind, Society, and Behavior

Background

The World Bank aimed to develop a toolkit that describes the latest findings in behavioral science so that researchers and designers in the development field have a systematic and accessible approach to use (World Bank, 2014).

Behavioral principles

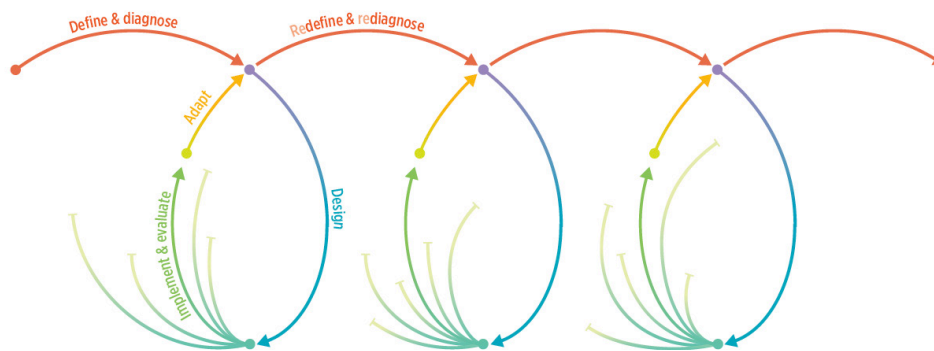
The World Bank's framework applies these three main principles:

- Thinking automatically: Much of our thinking is automatic, not deliberative. It is based on what effortlessly comes to mind.
- Thinking Socially: Humans are not autonomous thinkers or decision-makers, and we are deeply social animals. We have innate preferences for altruism, cooperation, and reciprocity, and we are strongly affected by the social norms and networks in our communities.
- Thinking with mental models: Individuals respond not to objective experiences but to mental representations of experience constructed from culturally-available mental models. Human decision making, therefore, is shaped by the powerful pull of both contextual cues and the past experiences of individuals and societies.

Design steps

This design approach can be described in these stages that repeat in cycles. They also emphasize investing more resources in the first half of the process:

- Define and Diagnose: Describe the behavior and the problem
- Design: Design an intervention that incorporates insights about the behavior and obstacles to it
- Implement & Evaluate: Experimenting with and testing different solutions to see what works
- Adapt: Revise solutions and re-define and re-diagnose the behavior



Source: World Bank (2014)

Environmental applications in GEF focal areas

- **Chemicals and Waste Management:** Behaviorally designed letters providing information on recycling and waste collection were delivered to 4,800 households and small businesses from the city of Trelew, in the Argentinean Patagonia (The Mind, Behavior, and Development Unit, 2019).
- **Land management and climate adaptation:** Examining the socio-ecological context of women's participation in forest conservation programs in Mexico (Walk a Mile in Her Shoes: eMBED Brief, 2018).

Strengths & weaknesses

- **Variety and strength of behavioral principles:** While the behavioral principles cover a range of insights, they do not cover more traditional strategies like information or incentives. It is also hard to distinguish how different strategies may be categorized.
- **Comprehensiveness of design steps:** The steps of this design process are generally strong, although it blends steps involving definition and diagnosis as well as implementation and evaluation. This reduces the amount of detail about the important actions to take for each of the above points.
- **Consideration of behavior change durability and evaluation:** This approach does not explicitly comment on the long-term durability of outcomes; however, it explicitly states that some interventions failed to produce long-term change. This provides some degree of perspective on the effectiveness of behavioral strategies.
- **Ability to work among diverse actors and scales of governance:** As an approach designed for the development context, it applies to a range of actors on a local, regional, and national scales.
- **Incorporates socio-ecological contextual analysis:** This approach emphasizes the importance of understanding the target actor's context as well as the social and structural obstacles they may face. It does not, however, provide a step-by-step guide or tools for understanding the greater system in which a behavior operates prior to exploring motivations and barriers.
- **Emphasis for design steps to be non-linear, adaptive, and iterative:** This approach has a strong emphasis on iterative and non-linear design, reminding designers to incorporate feedback from testing into future designs.
- **Consideration of ethics and power:** This approach specifically draws attention to the ways development designers make biased assumptions and how to acknowledge and address those when attempting to change behavior. It also describes the importance of understanding sensitive topics when doing research, and to be mindful of power dynamics in different contexts.

For more information on the World Bank's approach, explore the following resource:

World Bank. (2014). *World Development Report 2015: Mind, Society, and Behavior*. The World Bank. <https://doi.org/10.1596/978-1-4648-0342-0>

Ideas42

Background

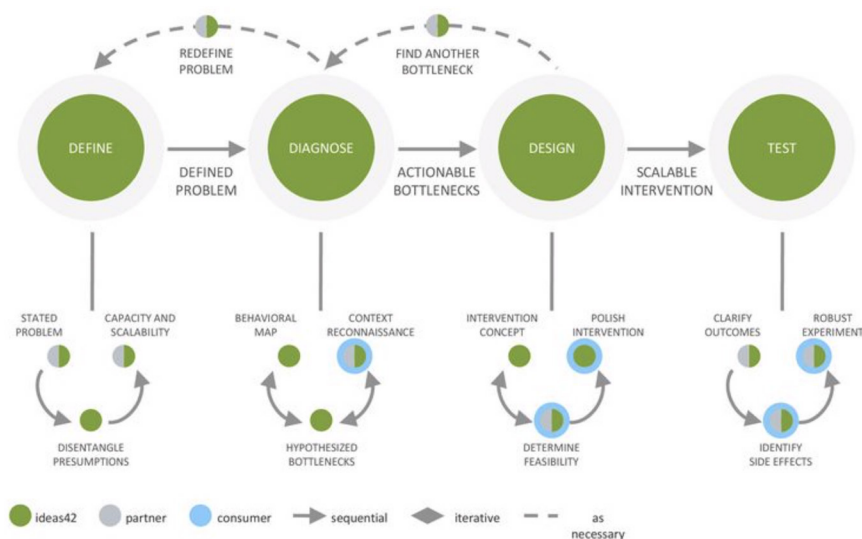
Ideas42 is a non-profit innovation firm that applies behavioral science for social impact. They have developed a standard approach and behavior change framework they use in all of their projects (Ideas42, n.d.).

Behavioral principles

- Choice overload: When faced with a huge range of options, many people fail to choose the best option or fail to choose altogether. Having more options often leads to less realized choices.
- Cognitive depletion & decision fatigue: Being tired and hungry can deplete our cognitive resources and significantly affect our decision-making.
- Hassle factors: Sometimes we do not act in accordance with our intentions because of seemingly minor inconveniences.
- Identity: Many of our choices are impacted by the perception we have of ourselves and our social roles.
- Limited attention: Our ability to pay attention to several things at once is much more limited than we might think.
- Loss aversion: People hate losses more than they love wins.
- Primacy bias: The bias toward the information that is presented first.
- Procrastination: We put off for tomorrow what we could (and often should) do today.
- Social norms: All of us are heavily influenced by our perception of what others are doing.
- Status quo bias: People prefer to stick with the status quo, even if other options are available.
- The availability heuristic: People judge probabilities based on how easily examples come to mind.
- The planning fallacy: People tend to be overly optimistic about their ability to finish tasks on time. They consistently believe the future will unfold as planned, and they rarely leave sufficient time to meet key deadlines.

Design steps (Tantia, 2017):

- Define: Identify the problem and intended outcome
- Diagnose: Hypothesize barriers to target behaviors relevant to your strategy
- Design: Develop strategies to address behavioral barriers
- Test: Testing your designed intervention through prototyping.
- Scale: Take your solution to scale.



Source: Datta, S., & Mullainathan, S. (2014)

Environmental applications in GEF focal areas:

- **Chemicals and Waste Management:** Before the tax went into effect, 82% of consumers shopping in sample stores in Chicago used at least one disposable bag per trip. Over the next year, the bag tax led to a decrease of 27.7 percentage points in the likelihood of using any disposable bags, despite the incentive being negligible (Homonoff et al., 2018).
- **Climate Change Mitigation and Adaptation:** Encouraging water conservation in Costa Rica: Two interventions—adding colored stickers focused on social comparison to utility bills and goal-setting postcards—each reduced water consumption by up to 5.6% (Datta et al., 2019).
- **Land Degradation:** Seaweed farmers learned to optimize growing practices when presented with summaries of results of experimental plots, with key dimensions highlighted (Faulstich-Hon et al., 2019).

Strengths & weaknesses

- **Variety and strength of behavioral principles:** The principles that Ideas42 describes on their website exemplify the ones they use in their work. These cover a range of specific behavioral insights, but they are not exhaustive in terms of all of the motivations and barriers to behavior (for example, they exclude incentives and information).
- **Comprehensiveness of design steps:** This approach includes a comprehensive and simple set of design steps ranging from identifying the problem to launching and scaling interventions.
- **Consideration of behavior change durability and evaluation:** The approach explicitly mentions different ways to test and evaluate effective solutions yet does not include any detail on monitoring over the long term.
- **Ability to work among diverse actors and scales of governance:** This approach has been applied in a variety of social impact contexts and among different actors at local, regional, and national scales.
- **Incorporates socio-ecological contextual analysis:** This approach emphasizes the importance of broader context—structural, economic, social, or psychological—for the successful development of interventions, yet does not offer explicit guidance for how to incorporate these insights. Ideas42 also has an Act to Adapt toolkit that specifies how different types of decisions (e.g., contextual, expert, political, structural) shape how an intervention is designed.

- **Emphasis for design steps to be non-linear, adaptive, and iterative:** This approach clearly describes the importance of iteration and feedback between the steps in its process.
- **Consideration of ethics and power:** This approach seeks to empower stakeholders and actors in their behavior change goals yet does not explicitly describe ways to mitigate power dynamics in each stage of the design process.

For more information on Idea42's approach, explore the following resources:

Ideas42. (n.d.). Ideas42. Retrieved June 17, 2020, from <https://www.ideas42.org>

Tantia, P. (2017). The New Science of Designing for Humans. Stanford Social Innovation Review. https://ssir.org/articles/entry/the_new_science_of_designing_for_humans

Appendix A:

Complete List of Reviewed Organizations and Their Behavior Change Approaches

| # | Organization/Authors | Approach/Tool Name |
|----|---|--|
| 1 | Behavioral Insights Team (Dolan et al., 2010) | MINDSPACE |
| 2 | Michie et al. (2011) | COM-B model, Behavior change wheel |
| 3 | Organization for Economic Co-operation and Development (OECD) | The Basic Toolkit for applied behavioral insights |
| 4 | Behavioral Insights Team | EAST |
| 5 | Ostrom (2007, 2009), Poteete et al. (2010) | Socio-ecological systems frame-work (SESF) |
| 6 | Rare's Center for Behavior and The Environment | Behavior change levers |
| 7 | European Union | Behavioral Insights applied to policy |
| 8 | eMBeD World Bank | Mind, Society and Behavior Report |
| 9 | United Nations (including UNEP) | Behavioral Insights 101 |
| 10 | Ideas42 | Ideas42 Principles |
| 11 | Consultative Group on International Agricultural Research (CGIAR) | Behavior change guide |
| 12 | National Oceanic and Atmospheric Administration (NOAA) | Risk Behavior and Communication: Best Practices |
| 13 | TRAFFIC | Persuasion toolkit to decrease the demand for illegal wildlife |
| 14 | Behavioral Science and Policy Association | Toolkit to strengthen the energy and environmental policy |
| 15 | Nordic council of ministers report | Nudging Pro-environmental Behavior |
| 16 | Behavioral Economics insights | Behavioral Economics Guide |
| 17 | Public health England | Toolkit for behavioral change for Weight management |
| 18 | Behavior Centered Design | BCD and Theory of Change |
| 19 | Behavioral Economics in Action at Rotman | A Practitioner's guide to nudging |
| 20 | Internal Revenue Service | Behavioral Insights Toolkit |
| 21 | Illinois Institute of Design | Brains, Behavior and Design toolkit |
| 22 | Common Cause Values Report | Framing Conservation |
| 23 | MDRC | SIMPLER |
| 24 | Bridgeable | A Guide to Using Behavioral Economics with Service Design |
| 25 | UNICEF | Demand for Health Services Field Guide |
| 26 | Unilever | Five levers of change for sustainable consumption |
| 27 | North American Association for Environmental Education | Influencing Conservation Action |
| 28 | World Wildlife Fund | Psychosocial Approach to Communication |
| 29 | Public Interest Research Centre | Framing nature toolkit |
| 30 | The University of Manchester | Change Points: A toolkit for designing interventions that unlock unsustainable practices |

References

- Cornish, A., Jamieson, J., Raubenheimer, D., & McGreevy, P. (2019). Applying the Behavioural Change Wheel to Encourage Higher Welfare Food Choices. *Animals*, 9(8), 524. <https://doi.org/10.3390/ani9080524>
- Datta, S., & Mullainathan, S. (2014). Behavioral Design: A New Approach to Development Policy. *Review of Income and Wealth*, 60(1), 7–35. <https://doi.org/10.1111/roiw.12093>
- Dolan, P., Hallsworth, M., Halpern, D., King, D., & Vlaev, I. (2010). MINDSPACE: Influencing behaviour for public policy [Monograph]. Institute of Government. <http://www.instituteforgovernment.org.uk/publications/>
- Faulstich-Hon, M., Sherman, E., Smith, G., & Ideas42. (2019). Act to Adapt: Behavioral Design for Climate Adaptation. https://www.ideas42.org/wp-content/uploads/2019/12/l42-1157_ClimateChangePaper_3.pdf
- Gorddard, R., Colloff, M. J., Wise, R. M., Ware, D., & Dunlop, M. (2016). Values, rules and knowledge: Adaptation as change in the decision context. *Environmental Science & Policy*, 57, 60–69. <https://doi.org/10.1016/j.envsci.2015.12.004>
- Green, K., Williamson, K., Park, T., & Reiner, C. (2019). Behavior Change for Nature: A Behavioral Science Toolkit for Practitioners. *Rare*. <https://behavior.rare.org/wp-content/uploads/2019/10/2019-Behavior-Change-for-Nature-Report-digital.pdf>
- Homonoff, T., Kao, L. S., Palmer, D., & Seybolt, C. (2018). Skipping The Bag. https://www.ideas42.org/wp-content/uploads/2018/09/Bag_Tax_Paper_final.pdf
- Ideas42. (n.d.). Principles. Ideas42. Retrieved June 17, 2020, from <https://www.ideas42.org/learn/principles/>
- Kropf, B., Schmid, E., Schönhart, M., & Mitter, H. (2020). Exploring farmers' behavior toward individual and collective measures of Western Corn Rootworm control—A case study in south-east Austria. *Journal of Environmental Management*, 264, 110431. <https://doi.org/10.1016/j.jenvman.2020.110431>
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, 66(1), 799–823. <https://doi.org/10.1146/annurev-psych-010213-115043>
- Michie, S., Atkins, L., & West, R. (2014). *The Behaviour Change Wheel: A Guide to Designing Interventions*. Silverback Publishing.
- Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science : IS*, 6, 42. <https://doi.org/10.1186/1748-5908-6-42>
- OECD. (2019). *Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit*. OECD Publishing. <https://doi.org/10.1787/9ea76a8f-en>
- OECD. (2017). *Behavioural Insights and Public Policy: Lessons from Around the World*. OECD Publishing. https://www.oecd-ilibrary.org/governance/behavioural-insights-and-public-policy_9789264270480-en
- Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, 11(6), 988–1010. <https://doi.org/10.3758/BF03196730>

- Palm-Forster, L. H., Ferraro, P. J., Janusch, N., Vossler, C. A., & Messer, K. D. (2019). Behavioral and Experimental Agri-Environmental Research: Methodological Challenges, Literature Gaps, and Recommendations. *Environmental and Resource Economics*, 73(3), 719–742. <https://doi.org/10.1007/s10640-019-00342-x>
- Rare. (2020). Levers of Behavior Change. Behavior Change for the Environment – Rare. <https://behavior.rare.org/behavioral-science-landing/>
- Reid, S., & Schmidt, R. (2018, September 18). A New Model for Integrating Behavioral Science and Design. *Behavioral Scientist*. <https://behavioralscientist.org/a-new-model-for-integrating-behavioral-science-and-design/>
- Rodríguez-Dowdell, N., Yepez-Zabala, I., Green, K., & Calderón-Villela, E. (2014). Pride for ARAs: A Guide to Reciprocal Water Agreements for People and Nature. *Rare*. <https://cdkn.org/wp-content/uploads/2013/01/RARE-ARA-Guide.pdf>
- Salzer, S. (2020, March 2). Behavioral Design 2020 and Beyond. *Medium*. <https://medium.com/behavior-design-hub/behavioral-design-2020-and-beyond-dc88a87f3b97>
- Shreedhar, G. (2020). Consumption and Biodiversity Conservation: A Mindspace Approach. In *Encyclopaedia of the UN Sustainable Development Goals*.
- Tantia, P. (2017). The New Science of Designing for Humans. *Stanford Social Innovation Review*. https://ssir.org/articles/entry/the_new_science_of_designing_for_humans
- The Behavioural Insights Team. (2012). EAST: Four simple ways to apply behavioural insights. https://www.bi.team/wp-content/uploads/2015/07/BIT-Publication-EAST_FA_WEB.pdf
- The Behavioural Insights Team. (2017). Evaluating the Nest Learning Thermostat. <https://www.bi.team/publications/evaluating-the-nest-learning-thermostat/>
- The Behavioural Insights Team. (2019). Increasing volunteer retention in West Java. <https://www.bi.team/publications/increasing-volunteer-retention-in-west-java/>
- The Behavioural Insights Team. (2020). Providing a substitute for single-use plastics in the Pacific. <https://www.bi.team/blogs/providing-a-substitute-for-single-use-plastics-in-the-pacific/>
- Thulin, E. (2020). Cooperative Behavior Adoption Guide: Applying Behavior-Centered Design to Solve Cooperative Challenges. *Rare*.
- Walk a mile in her shoes: EMBed brief. (2018). World Bank. <http://documents.worldbank.org/curated/en/955861557436108856/Walk-a-Mile-in-Her-Shoes>
- Wells, L., Salas, K. de, & Lewis, I. J. (2016). Behaviour Change Wheel Driven Normative Feedback in a Serious Game for Energy Conservation. *DiGRA/FDG*.
- Wendel, S. (2020). *Designing for Behavior Change: Applying Psychology and Behavioral Economics*. O'Reilly Media, Inc.
- World Bank. (2014). *World Development Report 2015: Mind, Society, and Behavior*. The World Bank. <https://doi.org/10.1596/978-1-4648-0342-0>

PART II:

Case Studies

Introduction

Many cases of behavior change, whether presented as verbal stories or research articles, focus specifically on interventions and their results. Yet, these lack the full picture of why and how a particular behavioral intervention works or the steps that implementers took to achieve a given behavior change result. In this section, we provide eight in-depth case studies that demonstrate the kinds of journeys that lead to successful (and sometimes unsuccessful) behavioral interventions. Case studies are instructive because of their detailed description of an intervention's effect but also because of what we can learn from the 'behind the scenes' of intervention development.

We chose the case studies in this section based on the following criteria:

- Real-world applicability: Cases feature field studies in real-world contexts
- Intervention-focus: Cases present solutions to a behavioral problem, not (just) to demonstrate behavioral insights
- Clarity: Cases demonstrate behavior change outcomes and why those outcomes did or did not occur is clear (ideally through a theory of change)
- Narrative journey: Cases tell a story about how the solution was developed
- Topic relevance: Cases are from one of the five GEF topic areas
- Strategy diversity: Cases feature interventions using a variety of behavioral principles
- Durability: Cases show the importance of monitoring effects over time (even if they lack that measurement in the results)
- Scalability: Cases show behavioral interventions across different actors and scales
- Replication: Cases have generalizable insights that could replicate in other contexts

For each case study, we present a short summary, the behavioral challenge and case background, the research and design of the intervention, the behavioral solution, the results, and the key lessons learned. After all of the case studies, we offer an analysis of what made them successful as well as where there is still room for improvement in behavior change work.



Figure 2: A map showing the eight case study locations, target behaviors, and target actors



Case Study 1: Water Consumption in Costa Rica

Summary

High rates of household water consumption were starting to deplete local water supplies in Belén, Costa Rica. To reduce water usage, the municipal government sent postcards to households comparing their water consumption to that of nearby neighbors. Postcards also provided tips for how residents could reduce their water usage and prompted them to plan how they would meet lower consumption levels. As a result, households reduced monthly water consumption by 4.5% (1.3 cubic meters) more than households that received the standard utility bill.

Challenge

Water consumption in Belén, Costa Rica (a town of about 21,000 residents) was 25% greater than the national average in 2010. It was estimated that if consumption stayed at that level, Belén could face significant water shortages by 2030. Reducing water consumption was thus a key policy priority for the Belén municipal administration.

Other interventions had not worked previously in Belén. A price increase in November 2012 had but limited impact even though water prices had increased by more than 100% (given the level of consumption of the average household). Fixed charges—for water use up to 20 cubic meters—increased by 70%. Despite these disincentives, by December 2012, the results showed that household water consumption had only decreased by 15% and rebounded to 5% higher than the November baseline by February 2013.

Meanwhile, the scope for price increases was limited by legal provisions stipulating that governments could not increase prices beyond what is needed to recover costs. Consequently, it was important to come with a simple and budget-light intervention to respond to the water urgency in Belén.

Research & Design

The team at Ideas42 developed their intervention using literature on behavioral interventions deployed in similar contexts, and from the findings of four focus groups run with diverse sets of Belén residents.

First, they found that residents generally believed in the importance of water conservation, but few of them believed that they individually should use less water. Residents justified their own water use based on Belén's climate, stating that it felt like a 'necessary evil.' As a result, the design team found that greater awareness may not be what was needed, but rather that personal appeals could be useful.

Second, residents were unaware of how much water they used. They focused on the cost of water instead of the amount of water used when talking about the information that they recalled from their water bills. Moreover, the water bills were hard to read and included information for both water and sanitation services, obscuring how much water individuals were actually using. Because of this, the team determined that making water use information more salient could be important.

Third, the water bills did not provide any comparison relative what reasonable or high amounts of water use were. Residents did not have an intuitive sense of what a small or large amount of water would be, and they did not know how to interpret cubic meters as a unit of measurement. As a result, the designers wanted to offer a relevant comparison for water usage—one in understandable units.

Finally, few participants were able to state concrete actions that would reduce their water usage. They could describe actions that used water but were unable to evaluate which of these used less or more water than others. The design team saw this as an opportunity to provide more guidance about specific water conservation behaviors.

Solution

The interventions were delivered as a one-time nudge through the mail. Households were selected according to postal routes, neighborhood, and average monthly water consumption; they were then randomized into three treatment groups and one control group. The first one, the “Neighborhood Comparison” treatment, took the form of a colored sticker pasted onto the monthly water bill. It communicated how a household’s water use compared to the median household’s water use in their neighborhood, including a smiley/frowny face to provide feedback about the household’s water use and water-saving tips. The second one, or “City Comparison” treatment, used similar stickers, the key difference being the use of city-wide, rather than neighborhood-wide, average water consumption as the reference point for the social comparison. The third treatment, a “Plan-making” intervention, consisted of a postcard that prompted people to record their water use, record a personal goal for water use, and check off specific actions that they could take to reduce their water consumption. This intervention applied salience, goal-setting, and implementation intentions to change behavior. The control group received a standard water bill with no additional information. Each treatment in this intervention took place in July 2014. The design team then looked at the August and September billing cycles to evaluate if any changes in water consumption had taken place.

Results

Overall, the residents who received the interventions conserved more water than those in the control group, but some treatments were better than others. The neighborhood comparison intervention reduced monthly water consumption 3.7-5.6%, while the city comparison intervention had no effect. The plan-making intervention was also effective at reducing water consumption between 3.4 and 5.6%. Overall, these interventions marked an average water reduction of 4.6%. The plan-making intervention appeared to be most effective for low consumption households, whereas the neighborhood comparison intervention appeared more effective for high consumption households. If all households in Belén received these interventions, they could save about 6,720 cubic meters of water each month. This translates to conserving 94,080 washing machine loads, 188,000 showers, or 222,000 dishwasher loads each month. Household savings were estimated to be between US\$2,600 and US\$5,200. The costs of the intervention materials themselves were also minimal, increasing the feasibility and impact of scaling the campaign to the whole municipality.

Lessons Learned

- Some behaviorally informed design strategies can translate across different settings and cultures. It is well-documented that the use of social comparisons is effective in changing energy consumption in a developed-country context. However, these interventions had not previously been shown to work in other contexts, particularly in low-income countries.
- This intervention was purposefully simple (two sets of stickers) because of resource constraints within the local government in Belén. This makes an intervention like this much more adaptable to other low-income country contexts, where complicated, software-reliant interventions may not prove a feasible solution.

- It is important to tailor behavioral interventions to the specific population and circumstance. As demonstrated in this case, focus groups with target users can help uncover key behavioral insights and guide the intervention's final design.

References and Related Studies

- Bernedo, M., Ferraro, P. J., & Price, M. (2014). The persistent impacts of norm-based messaging and their implications for water conservation. *Journal of Consumer Policy*, 37(3), 437-452.
- Datta, S., Miranda, J. J., Zoratto, L., Calvo-González, O., Darling, M., & Lorenzana, K. (2015). *A behavioral approach to water conservation: evidence from Costa Rica*. The World Bank. Retrieved from: <https://www.ideas42.org/wp-content/uploads/2015/04/Belen-Paper-Final.pdf>
- Miranda, J. J., Datta, S., & Zoratto, L. (2020). Saving Water with a Nudge (or Two): Evidence from Costa Rica on the Effectiveness and Limits of Low-Cost Behavioral Interventions on Water Use. *The World Bank Economic Review*, 34(2), 444-463.



Case study 2: Water consumption in Colombia

Summary

Managing water availability in urban spaces is a major challenge facing city governments and local councils. To address the potential water shortage in the city of Bogotá, Colombia, the municipal government implemented an awareness-raising strategy that initially backfired, increasing water consumption. After refining their approach, they discovered that leveraging principles of saliency, social rewards, social punishment, and education could more effectively change behavior and reduced water use by an astounding 13.8%. Moreover, the interventions established a new social norm of water conservation in the city that persisted ten years after the initial set of interventions ceased.

Challenge

In the 1970s, officials in Bogotá, Colombia, decided they needed to take action and address the city's almost-exponential increase in water demand in order to secure water for their growing population's future. Their first solution was infrastructural: a system called Chingaza that would involve a series of tunnels connecting nearby high-Andean water reservoirs to a water treatment plant and an emergency water reservoir in the city. They predicted this would satisfy water demand through to the early 2000s. But in 1997, a section of the tunnel that provided water to the water treatment plant partially collapsed. It was unclear how long it would take to repair the tunnel, and so it meant that only the emergency water reservoir, named San Raphael, was a viable water source. Experts recommended water rationing as the best short-term solution, yet the mayor, Antanas Mockus, was reluctant due to the political optics of such a decision. Instead, he wanted to build on his popular Citizenship Culture policy program.

The city declared a public emergency and started a communication campaign around a voluntary water savings program where everyone could cooperate to avoid a water crisis. The danger was that 70% of the city would be left without water if current water consumption was not reduced because water cuts would then be the city's only option. In the week after the campaign's launch, overall water consumption increased. The campaign had backfired. Some people even started stockpiling water to prepare for what they thought would be an imminent shortage. Even so, the city administration did not change their strategy until the lead academic of the administration's Citizenship Culture program, Paul Bromberg, investigated the problem.

Research & Design

Six days after the campaign launch, Bromberg conducted a quick, large-scale survey with residents to learn about why water consumption levels were increasing. They discovered that 96% of residents knew about the public emergency and why it had occurred. The results also showed that while 90% of people claimed to be trying to reduce water consumption, people were also mostly using strategies that were unlikely to make a difference, such as turning off the faucet while brushing their teeth. While the mayor, Mockus, did not have a clearly outlined theory of change, he followed social science closely, was open to trying things differently, and had a strong understanding of what could work for his city based on his Citizenship Culture program. This program emphasized building both moral and social norms for people to self-govern themselves and one another. Mockus further recognized

that Bogotans wanted something to be proud of after many years of war and violence in their city. He was two years into his administration when the crisis occurred and had already gained popularity and authority among the people, which would come to be one of his campaign's biggest strengths. There were also several shared cultural symbols and beliefs that were important to citizens, based on Bogotá being a largely conservative and religious population. The administration was further familiar with the importance of publicly recognizing local heroes in water conservation efforts, and with the fact that transparency from the government would be a key part of the efforts moving forward.

Solution

The city government launched an intensive three-month campaign. Part of it involved reminding people to conserve water at times when they were most likely to overuse it. To do so, the city distributed stickers that featured a picture of a statue of San Rafael, which was also the name of the emergency water reservoir that supported the city after the tunnel collapsed. Citizens were advised to place the stickers by the faucet that they used most frequently in the household, office, or at school. The stickers made the urgency to save water salient exactly when people needed to hear it most: when and where they habitually used water.

The city government also started entertainment campaigns that educated citizens about the most effective techniques of water conservation. Four thousand youth volunteers roamed the streets of Bogotá to inform people about the emergency and to teach them effective strategies to reduce water consumption. Catchy and memorable slogans carried the key messages of the campaign. The mayor appeared on TV showering with his wife and explaining that residents could save water by turning off the tap while soaping or by showering in pairs. Catholic priests were asked to invite their communities to join the cooperative effort. Newly-implemented technology-enabled daily reports of the city's water consumption were prominently published in the country's major newspapers. These sections were called Reloj de San Rafael (The San Rafael Clock) and were perceived as reliable and trustworthy sources of information that got cited on the radio and TV stations as well.

A third component of the intervention had the government publish information about who was cooperating and who was not. The CEO of the water company personally awarded households with outstanding water savings a poster of San Rafael and a title, "Here we follow a rational plan for using the precious liquid." These awards were visible in the media, and the government further spotlighted individuals' unique stories and water conservation strategies. The technicians working on the tunnel were also labeled as heroes in the city. On the day celebrating the completion of the tunnel's reconstruction, a second tunnel collapsed in the reservoir. At this time, Paul Bromberg had become mayor of the city, and Mockus, the previous mayor, had begun a formal presidential campaign. The city was concerned about ongoing cooperation and oversaturation from several months of media campaigns. Therefore, Bromberg added a new component to the intervention by creating city-imposed sanctions for "despilfarradores" (squanderers), those with the highest levels of overconsumption. While the sanctions were minor—squanderers had to participate in a water-saving workshop and had an extra day of water cuts—they were nevertheless effective because they targeted highly visible actors. Car-washing businesses, although collectively not a major source of water consumption, were the primary targets for this reason.

Results

The change in strategy helped to create a social norm of water conservation. By the eighth week of the campaign, citywide water savings had significantly exceeded the expectations and peaked at a reduction of 13.8%. Cooperation remained strong throughout the 290 days of the public emergency. Moreover, the reductions in water use persisted long after both tunnels were repaired. Per capita water usage remained lower than before the crisis for more than ten years. This suggests that the new social norm around conservation was established and persisted over time.

Lessons Learned

- Making the assumption that providing information is enough to change behavior can lead to unintended or even counterproductive effects.
- Combining multiple behavioral change principles like social rewards and punishments, salience, and messenger effects can be a useful approach to achieving substantial behavioral change.
- Building a community and making appeals to reciprocity can establish a social norm that results in a long-term change in behavior.

References and Related Studies

Arenas Gómez, J. C., Gutiérrez Sanín, F., Gutiérrez, M. T., Guzmán, T., & Pinto Ocampo, M. T. (2009). Politics and security in three Colombian cities. *Crisis States Research Centre*. <http://bibliotecadigital.udea.edu.co/handle/10495/10994>

Bernedo, M., Ferraro, P. J., & Price, M. (2014). The persistent impacts of norm-based messaging and their implications for water conservation. *Journal of Consumer Policy*, 37(3), 437-452

Guillot, J. (2019). Aplicación del enfoque de cultura ciudadana para motivar la reducción del consumo de agua a escala de ciudad: El caso de la crisis del agua en Bogotá en 1997. In *Lo mejor de las personas: Teoría, intervención y agenda de la cultura ciudadana*. Alcaldía de Medellín.

World Bank. (2015). *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank. doi: 10.1596/978-1-4648-0342-0.



Case Study 3: Increasing Farm Productivity in Kenya

Summary

Encouraging the adoption of new technologies to boost agricultural productivity and food security remains a challenge. To increase agricultural yields in Africa by the uptake of fertilizer, researchers developed a behaviorally informed program that addressed the issues of timing and impatience of the farmers. The Savings and Fertilizer Initiative (SAFI) program increased the fertilizer adoption by 13–14 percentage points.

Challenge

There have been consistently low agricultural yields in Africa, even as other regions (notably South Asia) have enjoyed dramatic increases in farm productivity. This is a vexing problem in agricultural development given the large number of people who directly depend on agriculture for both subsistence and income. One possible explanation for this discrepancy is the low adoption of inorganic fertilizer. If used correctly, fertilizer has the potential to dramatically increase yields and be a highly profitable investment. Moreover, in some ecosystems, the lack of available nutrients in the soil are unable to produce the yields needed to feed growing populations—this makes fertilizer a necessity.

The more conventional way to address this issue via economic policy has been to offer farmers a price subsidy on fertilizer. This strategy has been widely adopted in some African and South Asian governments, but large subsidies are financially demanding, typically regressive (benefiting the wealthiest farmers most), and often necessitate that the government's involvement in fertilizer distribution. Relying solely on price subsidies also assumes that the main barrier to increasing adoption is affordability. In other words, that the reason behind the low adoption of fertilizer is that farmers may not have the cash or credit readily available to buy it.

Research & Design

To develop the intervention, the researchers conducted a set of agricultural trials over six seasons in Busia District, Western Kenya. They began in July 2000, working in conjunction with International Child Support. The Savings and Fertilizer Initiative (SAFI) program was first piloted with minor variations over several seasons and at a very small scale with farmers participating in on-farm trials. The trials involved farmers working on their own farms in a region of Western Kenya where fertilizer use is low, and they revealed that when farmers use limited amounts of fertilizer, they can create profitable yields even without having to make other changes to their agricultural practices

Additionally, in November–December 2009, researchers asked 139 farmers in the same area about whether they had used fertilizer the past three seasons, and if so, when they had bought it. Despite the potential returns to applying limited quantities of fertilizer, only 40% of farmers in the sample reported ever having used fertilizer, and only 29% reported using it in at least one of the two growing seasons before the program. When asked why they did not use fertilizer, farmers rarely said fertilizer was unprofitable, unsuitable for their soil, or too risky; instead, they overwhelmingly replied that they wanted to use fertilizer but did not have the money to purchase it. At first this seemed difficult to take at face value, since fertilizer could be bought in small quantities (as small as one kilogram)

that offered annualized returns of at least 52%. Farmers could buy more fertilizer over time based on the profits from better harvests and eventually be able to fertilize a whole plot.

Apart from affordability, farmers also face issues of timing; there can be a mismatch between when fertilizers are typically offered, demanded, and needed for application, and when farmers have cash in hand to pay for it. At harvest time, when farmers have cash, they may not be motivated to buy fertilizer, and pre-purchasing it may be inconvenient. Later in the season, however, farmers may find that they do not have enough readily-available funds left to buy it, even though it is the right time to apply fertilizer. This mismatch suggests that just making fertilizers cheaper is insufficient. Instead, policies that encourage farmers to buy fertilizer when they have cash at their disposal (i.e., immediately after the harvest) could increase fertilizer adoption.

Some farmers may not have the money to buy fertilizer because they intend to save the funds but are unable to convert that intention to action later in the season. Depending on the season, 96-98% of those who used fertilizer had bought it just before applying it. Overall, depending on the season, only between 0.4% and 2% of farmers had bought fertilizer well in advance. There is also some anecdotal evidence that farmers do not follow through with their plans to buy fertilizer: 97.7% of farmers who participated in the demonstration plot program reported that they planned to use fertilizer in the following season, but only 36.4% of them actually followed through on their plans. This type of behavior can be explained by people's tendency towards being present-biased; people may plan to act rationally and in their interest in the future (e.g., intend to buy fertilizers), but when the future comes around, they behave impatiently (e.g., forgetting to buy fertilizer). As a result, even farmers who want to use fertilizer do not have the money to purchase it when the time comes to apply it later in the season.

Solution

In collaboration with the ICS, researchers designed an intervention to test if providing mechanisms to save harvest income for future fertilizer purchases could be effective in increasing usage. The researchers wanted to see how the SAFI program fared against other strategies to improve the uptake of fertilizer. Four different interventions were tested over two seasons with a set sample of farmers:

1. Basic SAFI: An ICS officer visited farmers immediately after the harvest and offered to sell them a voucher for fertilizer, at the regular price, with free delivery later in the season. In that moment, the farmer had to choose if they would participate in the program and how much fertilizer they wanted.
2. SAFI with ex-ante Choice of Timing: An ICS officer visited the farmers before the harvest and offered them the opportunity to decide when in the next season they wanted someone to ask them to participate in the SAFI program. An officer would then visit them and offer the fertilizer voucher as in the Basic SAFI.
3. Free Delivery Visit Later in the Season: Same as SAFI program, but ICS officers visited farmers 2-4 months after the harvest and offered to sell them fertilizer as a top dressing for the next crop.
4. Subsidy Later in Season: An ICS officer visited the farmers 2-4 months after the harvest and offered to sell them fertilizer, at a 50% subsidy, with free delivery.

Results

The SAFI program was very popular. In the two seasons that the basic SAFI was offered, the program increased fertilizer use by 14 and 18 percentage points, respectively (an increase of around 69%). The SAFI with ex-ante timing choice increased fertilizer use by 22 percentage points, and the subsidy intervention increased usage by 13 percentage points. These interventions were all more successful than the free delivery visit later in the season, which did not significantly affect usage. None of the interventions led to farmers using fertilizer in the following season—they were all only effective in the current season. This suggests that it was the lack of an available

commitment mechanism that was preventing farmers from purchasing and using fertilizer. Overall, the results suggest that offering farmers small, time-limited discounts on fertilizer may substantially increase usage without inducing overuse among farmers who are already using fertilizer.

Lessons Learned

- A cheap, small, and timely behavioral intervention via improved program design can be as powerful as a heavy subsidy.
- Designing an intervention based on the needs, preferences, and barriers of the farmers revealed through surveys and interviews was crucial for identifying the best components and options for the intervention. The SAFI program helped farmers overcome problems of mismatched timing and present bias, apart from affordability.
- Testing an intervention through pilots before launching at scale can provide important information about ways to make an intervention more effective
- Commitment devices can be powerful tools to help farmers make good investments.
- Behavioral interventions can be more effective if they simultaneously address the unintended negative effects of price subsidies, such as the environmental damage caused by fertilizer overuse.
- The effects of the intervention may disappear if the intervention is removed. Fertilizer adoption increased while advanced purchasing was offered, but once the program was removed, fertilizer usage went back to what it had been.

References and Related Studies

- Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *The Quarterly Journal of Economics* 121(2): 635-672.
- Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101(6): 2350-90.
- Duflo, Esther, Michael Kremer, Jonathan Robinson. 2008. "How High are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya." *American Economic Review: Papers and Proceedings* 98 (2): 482-488.
- James J., David Laibson, and Brigitte C. Madrian. 2011. "\$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans." *The Review of Economics and Statistics*, August 2011, 93(3): 748–763.



Case Study 4: Wild Meat Consumption in Brazil

Summary

Growing populations around the world lead to increasing meat consumption, which can have an impact on local wildlife. A group of researchers aimed to reduce wild meat consumption among households in the Brazilian Amazon town of Tapauá. They tried three different interventions that combined economic incentives with community outreach strategies and found that the outreach interventions were more effective than economic incentives alone, reducing wild meat consumption by 62%.

Challenge

As human populations grow, so does the demand for meat. In the Brazilian Amazon, growing populations are increasingly urban and can deplete surrounding wildlife populations. The wildlife trade in the Amazon is uniquely local and regional rather than international (as is the case with species in Asia and Africa). While this is positive for reducing wildlife exports, it also means that there is a smaller supply of meat to meet growing demands. Wildlife at risk from the consumption of wild meat includes various types of mammals, birds, and turtles, and tortoises. Researchers aimed to understand and change meat consumption behavior in the Brazilian Amazon town of Tapauá, which has a population of approximately 10,600 people.

Research & Design

The researchers first identified that their target behavior was not to stop wild meat consumption but to instead reduce it in a way that preserves cultural traditions and sustains wildlife populations. They also wanted to increase the consumption of chicken, pork, and fish as substitutes. They started with four focus groups to learn more about the motivations behind, and barriers to, consuming less wild meat. Three of the focus groups were with women, because they are those who most often buy and prepare the meat in Tapauá, and one focus group was with fishers because they could affect the supply of fish in the area. The researchers also conducted ten in-depth interviews with people who currently supply or consume wild meat so they could learn about what drives those decisions.

The findings of this preliminary research revealed that the poor taste and high prices of domesticated meat were barriers to their consumption relative to wild meat. People also liked the diversity and lower cost of wild meat and fish. They tended to eat wild mammals and birds during everyday meals, but turtles were a symbol of status and saved only for special events. In addition, the research team learned that chicken was the most available domestic meat, as compared to lamb, goat, and pork, and that there were gaps in people's understandings of wildlife ecology and population growth. Given the importance of price, the researchers wanted to learn more about what discounts would be most effective in encouraging domesticated over wild meat consumption. They randomly assigned different coupon amounts to 30 households and found that \$3-\$5 discount coupons were the most optimal for their intervention.

Solution

The researchers developed a theory of change for their intervention based on the Theory of Planned Behavior: an increased knowledge about ecology, more positive attitudes about wildlife, greater feelings of control over purchasing domestic meat, and a higher preferences for domestic meat was predicted to lead to an increased intention and commitment to consume more domestic meat and less wild meat.

They randomly selected and assigned 157 households to one of three treatments that included a combination of different behavioral strategies over three different periods of monitoring. These strategies were public information (e.g., visual media, mass media, giveaways, church visits, print media), community engagement strategies (e.g., door to door visits, commitments/pledges, cooking courses), and economic incentives (discount coupons for chicken in the intervention condition and coupons for cleaning products as the control condition). The treatments included: a community engagement group that received public information, community engagement strategies, and chicken discount coupons; a coupon group that received discount coupons for chicken and public information; and a control group that received public information and discount coupons for cleaning products.

Results

The researchers gathered self-report data during 15-20-minute visits they made to households 30 times over 60 days. During each visit, households reported the type and quantity of meat they consumed, how they got the meat, what the meat cost, and how many animals and which parts of the animals they consumed. Overall, households consumed fish and chicken the most as a result of the intervention. There was also an increase in knowledge about wildlife ecology across treatments, and those in the community engagement group increased their stated preference for fish, chicken, and beef. 92% of people in this group also made a public commitment to reduce wild meat consumption. The results showed that the \$3-5 chicken coupons increased chicken consumption and decreased fish consumption for households already consuming those meats but did not reduce wild meat consumption. In contrast, the community engagement treatment without any price incentives reduced wild meat consumption by 62%. The designers attributed this change in consumption due to the shifting perceptions around wild meat consumption from community engagement strategies. The provision of concrete tips for cooking domestic meat also likely had an effect.

Lessons Learned

- Material incentives are sometimes not enough to change behavior on their own and can lead to unintended outcomes
- It is important to be very specific about the target behavior, such as decreasing but not preventing wild meat consumption, or not including turtles in the category of wild meat
- Conducting research with the target actor reveals important insights and validates the underlying motivations and barriers needed to design a successful intervention
- Culture context can be a critical consideration for interventions in terms of what the actors will find appropriate and is likely to adopt
- Piloting different aspects of the intervention prior to its launch can increase its effectiveness

References

Chaves, W., Valle, D., Monroe, M., Wilkie, D., Sieving, K., & Sadowsky, B. (2017). Changing Wild Meat Consumption: An Experiment in the Central Amazon, Brazil. *Conservation Letters*. <https://doi.org/10.1111/conl.12391>



Case Study 5: Sustainable Farming in Mexico

Summary

The decline of indigenous agricultural systems has negatively affected Mexico's biodiversity, climate-resilience, nutrition, and public health. With the rise of commercial seeds, nitrogen fertilizers, and synthetic pesticides, many communities have abandoned ancestral farming methods such as Milpa, an inter-cropping system that enriches the soil, protects natural ecosystems, and produces high-nutrient foods. Rare and Centro de Investigación y Servicios Profesionales A.C. (CISPERP) developed a social marketing campaign to promote the re-uptake of traditional Milpa practices among farmers in the town of Tojtíc in 2018. After a year of the campaign and ongoing support from the Center for Behavior & the Environment, there was a significant increase in sustainable Milpa planting practices and consumption of Milpa products, leading to improved ecological, social, and nutritional outcomes for the Tojtíc community.

Challenge

Around the world, traditional agricultural systems and indigenous knowledge have been critical pillars of biodiversity conservation, as well as food and nutrition security for centuries. Milpa is an intercropping farming system consisting mainly of beans, maize, and squash, also known as "Las Tres Hermanas" or "The Three Sisters." These three crops form symbiotic relationships with each other and with other species, such as fruit trees and wild herbs. Traditionally, Mayans would preserve their native seeds to maintain the Milpa system alive. In this way, hundreds of variations of maize, squash, beans, and other crops would be preserved across generations.

However, with the recent rise in the use of commercial seeds and chemical fertilizers, many indigenous farming communities in Mexico have abandoned ancestral farming practices for new technologies promising higher yields at lower costs. Multinational corporations and government programs target and encourage farmers to adopt industrialized methods, guaranteeing higher profits with their use. They often do this by giving out free fertilizer and commercial crop seeds. Over the years, many Milpa farmers have started planting these commercial crops instead of their local seeds.

Replacing the Milpa system with industrialized agriculture further contributes to malnutrition and obesity, both of which are major public health challenges for Mexico. Along with this shift to commercialized foods, there has been a growing perception that Milpa is the food of the poor, or of the "Campesinos," casting a negative social image on the traditional dishes that use maize, beans, squash, chili, and wild herbs.

In the community of Tojtíc, in the Mexican state of Chiapas, the Milpa system has long been a part of the community's cultural heritage. The people of Tojtíc belong to the indigenous Mayan group, Tsotsil, and for centuries they have cultivated their own strands of maize, squash, beans, and other native crops. However, like many communities in the Mesoamerican region, Tojtíc was hit with a wave of campaigns and marketing efforts by national government programs in the early 2000s that sought to industrialize agriculture and drive economic growth. While most farmers in Tojtíc continued to value their native Milpa seeds, they began using commercial seeds and chemical fertilizers for their main production plots—moving their native seeds over to their smaller, household gardens. After adopting industrialized methods as their primary form of farming, the sustainable practices of intercropping and

natural nutrient management were increasingly lost or devalued. In response, the non-profit organization, Centro de Investigación y Servicios Profesionales A.C. (CISERP), began working with the Tsotsil farmers in Tojtíc to revitalize their traditional Milpa systems and promote agroecological practices.

Research & Design

In partnership with Rare, CISERP designed and launched a year-long social marketing campaign to address the social and environmental challenges facing the community of Tojtíc, with a focus on Milpa practices. Rare conducted capacity building training with local designers, developing a theory of behavior change, social marketing tactics, quantitative and qualitative research techniques, and a campaign designed to achieve social and ecological impacts that were community-led and human-centered.

In the first phase of the program, Rare helped CISERP develop a conceptual model that analyzed the systemic forces at play in Tojtíc and to determine which behavior, if changed, could have the biggest impact on biodiversity, food security, and ecosystem health. CISERP identified target actors of 324 Milpa farmers in the town of Tojtíc, and their campaign centered on promoting four target behaviors: compost organic waste (rather than burning it), reduce the use of chemical fertilizers, share and exchange native Milpa seeds with other farmers, and commit to a 5-year program throughout which these efforts would be sustained.

CISERP developed a theory of change based on a number of important insights from the results of 122 surveys and 20 in-depth interviews with farmers. First, farmers fondly remembered a time when their land was more biodiverse, and when they were not using as many chemicals. There is a strong connection to ancestral practices and ways of life, yet the introduction of chemicals made everything faster and easier, as well as creating a dependence on these products to get a consistently good harvest. There was little knowledge as to the effects of agrochemicals on farmers' health, their environment, and the community's economy, as well as on alternative options and support available for sustainable farming practices. There was also an opportunity to talk about the cost savings of planting Milpa, and to create conversation among farmers about the benefits of Milpa and share ideas about what was working. Farmers were very interested in participating in seed exchanges with others and already enjoyed spending their time doing a number of social activities. Finally, there was an interest and need for training and technical support, for celebrations of local seed varieties, and for long-term technical support to maintain new sustainable practices.

Solution

In collaboration with the community of Tojtíc, CISERP developed materials and activities to promote the target behaviors. The materials included a slogan, a comic book, a puppet show, and a mural, all in the native Tsotsil language, and all conveying that ancestral Milpa practices are something to be proud of and protect. These campaign materials were designed to be digestible and accessible, weaving the campaign message into the culture of the community. The puppet shows and comics were especially geared towards children, who were believed to then carry the message of pride home to their parents. Engaging the children was also a way of helping them develop an early appreciation for local and sustainable agriculture, something that was found to be largely missing because many young people in the region migrated to cities.

To train the farmers in the targeted agroecological practices, CISERP hosted hands-on capacity-building training sessions, called Field Schools, where farmers could gather to learn about sustainable practices like composting, intercropping, and nutrient management. To promote the conservation of local seeds and use of the Milpa crops, CISERP organized a seed fair for farmers to exchange their native seeds. Simultaneously, they also hosted a nutritional fair in which the women in the community opted-in to prepare dishes using ingredients from the Milpa and shared recipes with each other. Coupled with the social marketing materials, these fairs offered the community

a space to celebrate their native Milpa crops, participate in peer-learning, and preserve the agroecological practices unique to their culture.

Throughout the year-long campaign, Rare staff provided ongoing mentorship and support to CISERP to carry out these various activities and monitor the results.

Results

A year after launching their campaign, CISERP achieved promising results in the community of Tojtíc. The percentage of Milpa farmers using their own family's set of seeds and exchanging with other farmers jumped from 35 to 80%, meaning that 259 farmers were now participating in the practice of trading seeds and increasing the biodiversity of their farms. 65% of Milpa farmers committed to participating in a 5-year seed conservation program, supporting the durability of the behavioral shifts. CISERP utilized the campaign materials, such as the comic and puppet show, to reach other Tsotsil communities outside of Tojtíc and to spread the important message of conserving the Milpa.

There were lasting impacts beyond the immediate ones. The percent of Milpa farmers in Tojtíc who adopted the agroecological practices promoted in the campaign, such as composting organic waste, intercropping, and reducing the use of chemical fertilizers increased from 52% to 90%. There was a 95% decrease in the use of Glyphosate, the herbicide commonly used in Chiapas (the main ingredient in the "RoundUp" pesticide products). Additionally, 65% of the Milpa farmers had increased production and consumption of maize, beans, squash, and other crops associated with Milpa, thereby increasing the biodiversity of their farms and the nutritional value of their diets. An increase in the consumption of Milpa products also led to observable nutritional improvements. Over the course of the campaign, the average consumption of calories per day went from about 1900 to 1700—an effect attributed to a reduction in consumption of ultra-processed foods. For example, Pozol, which is a traditional refreshing drink made from fermented native corn, grew in popularity as a replacement to sugary drinks and sodas.

There was a complete eradication of the practice of burning residual Milpa waste, with 100% of the farmers transitioning to composting organic waste. Using the newly generated compost, CISERP trained a group of 16 farmers in Tojtíc in harvesting mushrooms for both self-consumption and sales. A total of 36 families in the community took part in this effort across four mushroom harvesting sites, each producing an average of 220 kg (485 pounds) of mushrooms over the course of 8 months—60% of which was distributed to the community for household consumption and 40% for sales.

Lessons Learned

- Understanding ancestral and cultural practices and traditions can provide a meaningful starting point for change, especially when there is already a need to revive and celebrate them in everyday life
- An intervention can help address multiple, related target behaviors with similar motivations and barriers
- Building opportunities for peer learning, descriptive norming, and capacity building can help to overcome major barriers to change
- A good intervention can affect many different types of indicators, such as environmental, social, and health indicators

References

Campañas por la conservación. (2018). Centro de Investigación y Servicios Profesionales A.C. (CISPERP).



Case Study 6: Reducing Overfishing in Indonesia

Summary

In the Dampier Strait of Indonesia, overfishing and the lack of rights to fishing areas were creating major challenges for coastal fishing communities. Leaders from 19 different fishing communities came together to co-create designated reserves and fishing areas (called Territorial Use Rights for Fishing, or TURFs) for each community and legalize their exclusive access and customary rights. These rights were also complemented by campaigns to adopt sustainable fishing behaviors, such as restricting the types of gear fishers used and fish they caught. As a result, the Dampier Strait network of 211,000 hectares is now the largest TURF + Reserve network in the world, and there has been a 44% increase in sustainable fishing behaviors.

Challenge

Overfishing is a challenge for many coastal fishing communities in some of the most biodiverse regions of the world. Having a reliable and abundant supply of fish underlies the livelihoods, economies, and food security of these communities. The Dampier Strait, a body of water between the Batanta and Salawati islands of Raja Ampat in West Papua, Indonesia, is one such place. There are 19 villages comprising ~6000 people along the strait that harvest fish there as well as small and large-scale fishers from other communities and countries. The challenge became to create a rights-based access system for fishing in Dampier Strait, coordinating this across villages, and to communicate the new rules to outsiders. While there was a history of customary rights for the Maya Tribe villages, there were no legal rights that allowed them to stop intruders from fishing in their waters and reduce the overfishing of limited resources for the communities that rely on them.

Research & Design

Rare's team sought to encourage behavior change from both local leaders, in creating this rights-based system, as well as fishers, to fish sustainability and manage the fishery through surveillance and reporting their catch. Fish buyers were also targeted to aid in documenting local catch. A number of relevant stakeholders were also involved in this solution, including religious leaders, media professionals, government units, and female leaders in the community.

The design team collected information from a representative sample of community members from 19 villages in Dampier Strait. They conducted interviews, surveys, and observations with village governments, customary leaders, and fishers as well as staff at the District Office for Fisheries and Marine Affairs and the District Technical Implementing Unit of Raja Ampat.

For local fishers, a number of key barriers and motivations emerged. Fishers knew that they were competing with other fishers, but they were hesitant to say anything to them for fear of creating conflict. Trust was a large barrier. Over 60% of fishers in Dampier Strait distrusted fishers from outside their village due to these interactions, and the average trust in fellow villagers was at an average of 55%. The trust in the local government was also about 50%.

Knowledge was also a major barrier. There were misperceptions and denial about the impacts of overfishing on fish stocks, where 40% of households perceived no link between human activity and fish populations. There was no prior knowledge about TURF systems, which also demonstrated to the campaign team that any print media they had provided on TURFs in the past was not working and that they needed to diversify their use of information platforms. There were also low levels of self-efficacy in making decisions that could make a difference as well as in social support for change.

Besides barriers, the campaign team identified a number of motivations for change. Formal and customary (adat) leaders were very influential in setting expectations, values, and norms for behavior. There was also a clear interest in having fish populations that could sustain for future generations. For village leaders, the campaign team found that there was a big motivation for them to be seen as responsible and wise in supporting the interests of their villagers. For buyers, the biggest motivation was to be able to predict future revenue from having a record of their transactions.

Solution

The design team launched a 12-month behavior change campaign targeting fishers, village leaders, and fish buyers. In order to design a rights-based system and appeal to village leaders, Rare and the provincial government hosted a series of workshops where the village, customary, and religious leaders of 19 communities drew maps of what they considered to be their fishing grounds—working together to resolve conflicts and identify the boundaries for each community’s coastal waters. This resulted in a single map of 21 access areas, or Territorial Use Rights for Fishing (TURFs), one for each of the 19 villages as well as several common fishing areas that villages would share. This provided the foundation for the rights-based system that, beyond granting exclusive rights to local fishers, also outlined the types of gear that could be used and the fish that could be caught in the fishery. The final designations were no-take zones, or marine reserves, that would be off-limits to fishers and help the fish populations grow.

To maintain the sustainable fishing behaviors within the newly designated areas, the campaign also used several behavioral strategies designed in collaboration with the community: facilitating the registration of fishers’ boats; highlighting key messengers in the community who supported the new rights-based system; organizing pledges in the community to uphold the fishing regulations; launching community surveillance of coastal waters and providing fuel for those doing the surveillance; controlling fishing inputs and outputs; offering training sessions on sustainable fishing; and integrating an app called OurFish into fishing operations to track fish catch and sales. There were a number of social marketing materials that complemented these activities, including calendars that reinforced fishing rules and campaign messages in each season, festivals to declare the TURF areas, puppet shows with children, a fish game that demonstrated the importance of cooperation and rights-based fishing on fish populations, and smaller giveaway items such as t-shirts, bags, and cushions that helped to create additional campaign visibility.

Results

The resulting network, officially designated a TURF + Reserve network at the national level, encompasses 211,000 hectares in Dampier Strait across 19 villages, making it the largest comprehensive TURF + Reserve network in the world. Between 2017-2019, there were large changes in beliefs and behavior. There was a 44.5% increase in compliance with fishing regulations, an 82% increase in knowledge of rights-based fishing areas, a 29% increase in trust in local villagers, 25% increase in trust in local governments, and a 15% increase in trust of other villagers to comply with regulations. There was also, on average, a 34% increase in participation for the development, management, and surveillance of the designated fishing areas. The environmental outcomes were also significant. Between 2016-2018, there was an 11% increase in coral cover and a 71% increase in fish biomass.

Lessons Learned

- Co-creating solutions and engaging many actors to help scale interventions provide robust and durable solutions that will be supported and reinforced by many different people
- Incorporating customary, religious, or traditional values can help to legitimize solutions among a range of actors
- Addressing different motivations through a range of strategies and behavioral principles makes an intervention increasingly effective
- Being intentional about the benefits and barriers for each target actor leads to more effective messaging, such as boosting self-efficacy and perceptions of leadership

References

Kushardanto, H., Rukma, A., Jakub R., Tjandra, K., Susanto, HA., Arum, GS., de la Rosa, E., Suherfian, W. & Campbell, SJ. (2019). Final Report: Improving Food Security and Livelihoods Through Rights-Based Management of Coastal Fisheries in West Papua. USAID SEA Project, 54pp.

Kushardanto H, Tjandra K, Jakub R, Campbell SJ. (2018d). Progress Report: Social Marketing Campaign for Mayalibit Bay and Dampier Strait. USAID-SEA Project, 18pp.

Kushardanto H, Tjandra K, Jakub R, Wibowo, JT, Campbell SJ. (2018b). Activity Report: TURF Social Marketing Campaign Strategies. USAID-SEA Project, 15pp.



Case Study 7: Silvopasture in Colombia

Summary

The adoption of silvopastoral systems (SPS) by Colombian cattle ranchers is necessary for both the sustainability of the country and the livelihoods of small and medium-scale farmers. Beliefs about the use of trees in grazing pastures and a lack of access to resources have limited the implementation of sustainable grazing methods. Project designers developed a multi-pronged approach to encourage SPS adoption in five regions of Colombia through improving access to financial and technical resources, increasing peer-to-peer knowledge transfers, enhancing the productivity and sustainability of ranches, and raising the standard of living for farmers and their families.

Challenge

The land area devoted to the cattle ranching in Colombia has increased dramatically, rising from 14.6 to 38 million hectares in the last 50 years. Currently, 38% of the total land surface and 89% of all agricultural land in the country are set aside for raising cattle. Not only is the cattle ranching process responsible for land degradation, it also employs methods that are both expensive and inefficient for small and medium-scale farmers.

Deforestation is a common practice among ranching communities in Colombia. As land becomes unsuitable for grazing, farmers clear the Andean and Amazon Basin forests and convert it into new pastures. This cycle continues, forcing farmers to expand into primary forests to make up for the degraded land. The clearing of land and the expansion of current cattle ranching practices both threaten tropical rainforests, biodiversity, and the carbon emissions pathway of the country.

As opposed to much larger, industrial, bovine farms, many of these local enterprises raise fewer than 50 cattle at a time. This limits a farmer's ability to adopt new technologies that would otherwise help their productivity and develop the sector. Regardless of a farm's size, the cattle ranching industry is an important facet of the Colombian economy, comprising 3.5% of the country's total GDP and at least 27% of the GDP for agricultural activities. However, these farms are typically located in communities suffering from high poverty levels, violence, extreme income disparities, and inequitable land ownership.

Silvopastoral systems (SPS) provide a more sustainable alternative to current ranching practices, one that would allow farmers to support the same number of cattle on less land. These systems involve combining a variety of vegetation, such as trees, shrubs, and herbaceous plants, for grazing. The benefits of this method of ranching also include increased local biodiversity, carbon sequestration, improved water quality, elevated productivity, and lower operational costs. Still, despite their benefits, silvopastoral systems have, historically, only had limited adoption rates among Colombian cattle ranchers.

Research & Design

This project was developed based on the lessons learned from previous silvopastoral programs in the region and from a variety of social and institutional assessments. Between 2002 and 2008, the GEF implemented a project to promote silvopastoral practices in Costa Rica, Nicaragua, and Colombia. Although this project indicated that the productivity and sustainability of cattle ranching in Colombia could be improved by silvopasture, its adoption within the program was limited by high costs, the complexity of methods, and a lack of knowledge by farmers. Furthermore, the program lacked the necessary incentives for adoption; many of the benefits of silvopasture were perceived by farmers as external to their operations. The coexistence of trees on grazing lands is also seen as problematic; farmers worry about the competition between the pasture for resources and the potential of falling branches that could injure cattle. This project illustrated that, despite comparable implementation across farms, silvopastoral practices did not always provide the same profitability to farmers nor contribute equally to conservation outcomes. Learning from this, planners chose certain strategies to improve the adoption of sustainable cattle ranching methods.

Program designers employed a variety of survey and assessment measures to better understand the social and political risks of the project, the socio-economic situation of ranchers, and the barriers to adoption for small and medium-sized farm participation. To further connect with the target actor, project partners completed surveys of sample farms, held workshops within the project regions, and conducted interviews with producer associations, environmental organizations, educational institutions, banks, and a variety of NGOs. In addition, assessments of larger social and institutional barriers to program participation were conducted to inform the strategy in different regions of Colombia. Once the project plan was designed, local partners led five regional workshops with farmers and extension agents to clarify the program components.

Solution

Using the lessons learned from past cattle ranching sustainability programs as well as the recent surveys and assessments, the project designers implemented a multipronged approach to increase the adoption of silvopastoral systems (SPS) by small to medium scale cattle farmers in five regions of Colombia. The solution included four distinct components: improving the productivity of farms through SPS, increasing connectivity in farms via payment for ecosystem services, strengthening monitoring and evaluation and related subsector institutions, and project management. Overall, the program could be broken into two parts: financial and technical assistance.

In providing financial assistance, cattle ranchers were given increased access to loans and microfinance investments, as well as adequate payment for the ecosystem services protected by SPS. These financial incentives also helped ranchers acquire new agricultural technologies and seeds that were previously unattainable in the regions.

Increased technical assistance was achieved through a variety of methods. One element was the use of project agents who provided SPS expertise via training sessions and local assistance centers for easy access by farmers. Another way this was facilitated was through peer-to-peer information exchanges between the farmers who were utilizing SPS on their land and those who were not utilizing SPS. This harnessed the power of social influence to encourage farmers to adopt these practices. Lastly, the program set up a silvopastoral certification process for farms, making it both accessible, valuable, and simple to achieve SPS designation.

More broadly, the program was advertised through radio stations, bulletins, and news sources, further increasing a normative shift towards SPS. To ensure community satisfaction, local participation was monitored, and feedback was collected through community complaint boxes set up in easily accessible locations.

Results

Prior to the implementation of the project components, a baseline assessment of productivity, environmental health, and socioeconomic level was conducted on each participating farm to determine the impact of the project post-intervention.

Since the project's inception, 4,100 family farms in five different regions of Colombia have adopted silvopastoral techniques in cattle ranching. Over 32,000 hectares of land were converted to silvopasture, and incomes increased by \$523 per hectare, per year. The milk productivity of farms also increased by 36.2% since conversion. Overall, participating farmers have experienced increased pride and investment in their work as cattle ranchers.

In addition to the localized benefits of the program, the project led to ecosystem-level improvements. Participating farms planted over 2.6 million trees, of 80 distinct species, that captured more than 1.2 million tons of carbon. Moreover, 21,000 hectares of key habitats for biodiversity were conserved and improved through the adoption of silvopastoral techniques.

Lessons Learned

- Building on the experiences and results of prior projects can provide a starting point for intervention design
- The use of common language in disseminating information on different SPS schemes is necessary for adoption.
- Providing a baseline incentive for farmers to protect ecosystem services counteracts the need for farmers to clear the land for grazing.
- The technical assistance provided through experienced and trusted extension agents has proven to be the most effective component of the intervention

References

Trees and cows offer path to recovery in Colombia. (2019, August 7). Global Environment Facility. <https://www.thegef.org/news/trees-and-cows-offer-path-recovery-colombia>

World Bank. (2009). Project appraisal document on a proposed grant from the Global Environment Facility to the Colombian Cattle Ranching Association for a mainstreaming sustainable cattle ranching project.



Case study 8: Increasing Rhino Monitoring in Namibia

Summary

Despite numerous conservation efforts across Africa, black rhinos remain one of the most critically endangered species in the world, and traditional fortress models of conservation have resulted in a lack of engagement and investment in the people living near parks and reserves. Community leaders in Namibia identified opportunities to reduce poaching by introducing a Rhino Rangers program where local people were chosen by their communities to track, monitor, and protect rhinos in their conservancy, as well as to participate in a network of other rangers to help capacity-building efforts, training, and support. As a result, poaching declined 83% during the first five years of the program, with no further poaching incidents reported to date.

Challenge

Poaching has been a threat to black rhino populations for decades, with the loss of 97% of the population from 1970-1990. Rhino populations experienced a brief recovery, but growing demand for rhino horn threatens rhinos with extinction. The illegal wildlife trade within and between Africa and Asia is a pervasive problem, and tracking and enforcement are difficult due to the illicit nature of poaching behavior and its complex trade networks. The highest numbers of rhino poachings have occurred in South Africa, Namibia, Zimbabwe, Botswana, and Kenya. Fortunately, Namibia has a record of successful community-based conservation practices. In 2011, local community leaders and guards felt a need to make a change to their practices to try and reduce poaching pressure on rhino populations. Historically, military-style law enforcement had been the most common strategy for protecting rhinos and has had little impact. Field conservationists from the existing Rhino Custodians Program in the area noticed that existing efforts tended to focus too narrowly on poaching and labeling poachers as criminals, rather than on the larger system that was causing people to poach rhinos in the first place. Additionally, models of fortress conservation in local parks were perceived as prioritizing the needs of wildlife over people—contributing to local communities' lack of interest in participating in conservation.

Research & Design

Based on what they knew about the communities' perception of conservation, staff from the Rhino Custodians Program started by developing an understanding of how people valued rhinos as well as who got to make decisions about rhino conservation. Their goal was to shift conservation from a top-down model to a community-based approach, with local ownership and multi-stakeholder engagement at local, regional, and national levels.

Leading up to the launch of the Rhino Rangers program in 2012, there was a detailed 2-year planning process. In 2010, there was a multi-stakeholder rhino security workshop with rhino experts across several countries in Africa who shared their experiences. Save the Rhino Trust helped to organize NGO support for increased rhino monitoring and conducted a needs assessment for rhino patrolling with current guards. The Minnesota Zoo and The Nature Conservancy also helped by providing seed funding for the Rhino Ranger program. In the summer of 2012, nine conservancies selected the first rhino rangers that would become part of the first cohort.

The Rhino Ranger program was built upon Herold Lasswell's value concept to help bolster engagement in the Rhino Ranger program. These values include decision making power, knowledge, wealth, well-being, skill, affection, respect, and rectitude, and the Rhino Ranger program developed a theory of change based on how these values were perceived and experienced in local communities. They identified that people needed to be engaged and empowered to protect rhinos; have the tools, skills, and motivation to perform conservation actions that also improved their well-being; wanted to participate in efforts coordinated through a local institution with government endorsement; and had access to income-generating mechanisms and community benefits that were linked to this local institution. As a result, they expected that the community would protect and invest in rhino protection; the quality and quantity of rhino monitoring would increase; the relationship with people and rhinos would improve; the rangers would gain respect; poaching would be less tolerated; and ultimately, the rhino population would grow.

Solution

The resulting solution was a complex set of strategies and tactics involving a nested structure of different actors at different scales: local rhino custodians, regional community rhino custodianship, and the national ministry of environment and tourism. The first initiative was called the Conservancy Rhino Ranger Incentive Program. This involved a group of rhino rangers who were chosen by and accountable to their local communities to monitor and generate income from rhinos on their conservancy land. They also were a part of a support group where they received a comprehensive set of incentives, training, and tools to help them perform their work, and were part of a broader strategy to support rhino monitoring and, in turn, reduce poaching.

The campaign logo and uniforms helped to build and strengthen a sense of identity, belonging, and pride for rangers. Hero jackets were also awarded by respected figures in the community after five years of service, in addition to display pins, monitoring performance stars, and the ranger's local affiliation. There were even custom-made memorial tombstones for rangers who passed away that honored their life of service. A three-part training curriculum and team building sessions provided the information and peer support for doing the job effectively. During the basic training, rangers read, pledged, and signed the rhino ranger honor code in front of their peers and received a bracelet to signal their commitment. Ranger-selected patrol food, monitoring technology, notebooks, diaries, logbooks, and camping gear provided the essential equipment for rangers to perform their duties without having to purchase these themselves for the two to three weeks a month they were in the field. In the field, rangers also had laminated rhino identification cards to strengthen their relationships with the rhinos in their conservancy. There were additional extrinsic and intrinsic incentives: monthly performance bonus payments for exceptional quality and quantity of reporting as well as awards posted in public operations room for the best photo, best sighting, and best overall patrol. Each element contributes to the ongoing program's theory of change, and together, they serve to increase the rangers' dedication to their role as rhino custodians, make their jobs easier, as well as strengthen the relationship between people and rhinos.

Results

Results of a pre-post assessment from the program show that poaching declined by 83% from 2012–2017. There were no poaching incidents from August 2017 to May 2019. A survey with approximately 300 farmers in 2017 showed that there were positive attitudes towards the Rhino Rangers and interest in having more in the community. Additionally, where community members previously shared no prior information with the rhino rangers about potential poachers, between 2017 and 2018, there was an increase of public information given to the rangers that allowed local enforcement in 11 of 16 cases to prevent poaching in the area. Finally, the number of rhino sightings from when the rhino rangers started in 2014 until 2018 increased from 79 to 918.

Lessons Learned

- Decision making driven by local stakeholders creates both more durable behaviors over time as well as reinforcing ones supported by a range of actors
- Providing basic needs and requirements to do the target behaviors are important (gear, food, equipment, transport)
- Collaboration built through a shared identity and purpose can drive behavior change (reinforced by branding, logos, slogans)
- Open communication and engagement are key to achieving buy-in from a diverse set of stakeholders, and involving stakeholders at multiple scales reinforces target behaviors
- Developing a core set of outcomes that are selected, measured, and monitored is helpful for evaluating success
- Addressing a wide range of motivations and barriers can make an intervention more successful

References

Muntifering, J. R. (2019). Large-scale black rhino conservation in north-west Namibia: A review of the Conservancy Rhino Ranger Incentive Programme. Venture Publications. https://issuu.com/travelnewsnamibia/docs/srt_booklit_final_issuu

Conclusion: Case Study Recommendations

There are a number of recommendations and lessons learned that emerged from the eight case studies.

Base the intervention on the needs, experiences, and socio-ecological context of the target actors.

All of the case studies grounded their interventions in research conducted with the target actor. This essential step uncovers insights about their motivations and barriers that then guides which behavioral strategies will be most effective at changing behavior. For several target groups, boosting their sense of self-efficacy and celebrating natural resources created change. For others, there were significant financial, time, or social barriers that needed to be overcome. These cases further demonstrate how those designing interventions need to put aside their assumptions about what might work and instead use the information they have been given, even if it is surprising or something with which they disagree.

Moreover, this highlights the need to understand the broader socio-ecological context, structural barriers, and systems for designing behavioral interventions. Only about half of the cases here had strong processes for doing so. For example, in Indonesia, the role of customary leaders and the principles of adat were key in developing a locally-recognized and legitimate designated fishing areas. In Brazil, wild meat consumption was an important part of cultural traditions, and certain wild animals were consumed at certain times. Women were also the primary food buyers in the household and responsible for preparing meals. In Mexico, planting Milpa had a significant ancestral legacy and value in the community, and it was important to understand how this had been replaced by industrial agriculture. In Namibia, a legacy of fortress conservation tactics forced local conservationists to re-envision conservation by examining existing relationships between people, wildlife, and local decision-making powers to effectively protect rhinos. All of these insights were incorporated into the design of the behavior change interventions.

Combining multiple behavior change strategies can be more effective than a single strategy.

All case studies applied multiple types of behavior change strategies and appealed to three or more motivations and barriers. Indeed, the few cases where there was an intervention that only used one strategy, and particularly information and incentives, the intervention backfired. For example, launching an awareness-raising campaign on water consumption in Bogota initially increased water consumption. It is important to know the limits of any given strategy and once again base the intervention in what the data say about the target actors, rather than what designers think makes the most sense to implement. The most common strategies were informational and social and used together. These included providing tips, training sessions, media, or feedback as well as and changing social expectations, creating opportunities for peer learning, and making behavior more observable. These strategies were not random or interchangeable; instead, they were uniquely chosen due to their ability to address the target actors' motivations and barriers. At the same time, just because cases use multiple strategies does not mean they need to be overly complex or expensive. For example, in Costa Rica, postcards featuring a combination of stickers that conveyed social and how-to information were effective in changing behavior. In Kenya, the use of a small and well-timed incentive could make the difference between applying fertilizer or not in a given season.

It is crucial to have a robust monitoring and evaluation framework for assessing behavior change.

Most case studies had a theory of change that linked to core behavioral, social, and/or environmental outcomes, yet these could have been improved. The case in Mexico was one of the strongest examples of accounting for social, environmental, and health outcomes. None of the case studies identified a specific psycho-social theory of change. Very few use a robust evaluation framework such as randomized control trials, and most rely on pre-post measurements of change. Very few of these cases had controls in their intervention design, likely due to the nature of these cases being field studies. Additionally, a small minority of case studies had explicit plans to monitor behavior change beyond the end of the intervention, which puts into question their long-term durability. Overall, even for the strongest case studies in this set, there is room for improvement in their monitoring and evaluation framework.

Engaging a range of stakeholders who have different roles and operate at different scales can enhance the intervention.

A number of these case studies consider several stakeholders either as part of the target actor group or as other influential actors in effecting behavior change. A single campaign can address the needs of multiple actors if there has been sufficient research and system mapping of their behaviors, motivations, and barriers. In the case of the Rhino Rangers program in Namibia and the sustainable fishing regulations in Indonesia, involving actors at multiple scales helped to reinforce the intervention and hold actors accountable to change. This also has implications for the durability of the intervention, since there are more individuals from local to national scales who are invested in the solutions over the long-term. Beyond working with actors at different scales, several cases were mindful of involving actors who held different roles in the community. For example, the Brazil case talked to women and fishers because of their impacts on the behavioral problem. In the Bogota water crisis and Tojtic cases, children were key messengers and educators of local citizens regarding the benefits of conserving water and farming sustainably.

It is necessary to have a very clear and specific target behavior and actor.

All of the cases demonstrate the importance of identifying a specific target behavior and actor (or several) when designing an intervention. For example, in Brazil, the researchers made it clear that they were aiming to decrease, not stop, meat consumption, all while increasing domestic meat consumption. They also specified that wild meat did not include turtles due to their unique cultural significance and rare consumption relative to other wild species. Their target actors were also narrowed to households, but particularly women who were the ones who bought and prepared meat for their family. Similarly, the case study in Indonesia had three different target actors (fishers, local leaders, fish buyers) and a number of different target behaviors for each that they addressed during the campaign. In Costa Rica, it was much simpler: households reducing water consumption. No matter the number of behaviors and actors, it is important that they are clearly specified in order to conduct the most accurate research and to design effective solutions.

Testing components of the intervention before going to scale can help designers create more effective solutions.

Several cases incorporated steps to test elements of their interventions and made revisions to their solution design. For example, in Kenya, the research team conducted trials over six seasons in order to learn more about the effects of offering incentives for fertilizer at different times. The results of these trials allowed them to design better-informed options for their intervention at scale. In Brazil, the research team did a small experiment about the level of incentive (i.e., a discount on chicken) that would be appropriate and increase the likelihood of purchasing

domestic meat. They used their findings in the design of their eventual intervention, and were then more confident that their findings were due to the incentive being offered and not the 'amount' of the incentive. In both Mexico and Indonesia, campaign materials were shared or even codesigned with local community members to incorporate their feedback on the brand and the type of messaging that would resonate best with the whole community. These instances show the importance of validating assumptions about the intervention before launching at scale and increase the likelihood of success.

PART III:

Selecting, Measuring, Monitoring, and Evaluating Behavior Change Indicators

Introduction

Not only is the monitoring and evaluation of program indicators critical for accountability, it is also required if implementers are to improve program development, provide rapid data for decision making, and to allow for the generalization of findings to new and varying contexts. Despite the large majority of environmental programs implicitly requiring behavior change, many programs fail to include behavioral outputs as key indicators. It is even rarer for programs to systematically evaluate psychological and social indicators. This leaves programs with both poor metrics of success and an even poorer understanding of the underlying reasons for why a program achieved the level of success it did, and whether it will generalize to novel contexts.

This section is broken into four components. In the **first component**, we provide guidance on how to select behavioral, psychological, and social indicators. Since different programs tackle different challenges in different contexts, we cannot make general statements about the specific indicators that should be included. Instead, we provide guidance on how indicator selection can be based on a program's psycho-social theory of change. This, for example, includes the explicit representation of how program activities affect the psychological and social characteristics of a community, which then affects behavior and achieves program outcomes.

The **second component** focuses on how psycho-social and behavioral indicators can be measured. This includes a discussion of direct outcome observations, an option available for some 'publicly accessible' behaviors; proxy measures, which can indirectly assess an indicator; and self-reports, which are often necessary to measure psychological and social indicators.

The **third and fourth components** discuss how the monitoring and evaluation of psycho-social and behavioral indicators can improve program development. The monitoring of key psycho-social indicators allows not only for adaptive management, where program design is improved over time (based on insights into the program's mechanism) but also for what we term dynamic programming, where live programmatic decisions about phase transitions, expansion, or termination can be made based on real-time monitoring of psychological and social states within a target actor. Moving to evaluation, the quantification of program outcomes, including changes in target behaviors, allows program designers to assess program efficacy and effectiveness in order to make informed decisions over the scaling of a program. Finally, we conclude with a discussion of how the integration of monitoring data of psycho-social indicators with outcome evaluation allows for informed decisions about the generalizability of programs to novel contexts.

Indicator Selection

Selecting indicators is a critical component for developing the monitoring and evaluation (M&E) plan of any program—behavioral change efforts are no exception. Traditional M&E frameworks in the environmental field tend to track program delivery components for the monitoring and evaluation of environmental impact. Tracking program delivery and environmental impact is critical for successful M&E. In this section, we aim to expand from these two types of indicators and provide guidance on additional monitoring and evaluation components that are relevant to programs targeting behavior change.

A Psycho-Social Theory of Change

Emerging in the mid-90s, *Theories of Change* (ToC) have become the dominant paradigm for program planning and monitoring (Coryn et al., 2010). At its core, a program's ToC is a directional flow diagram that illustrates the relationship between program elements, intermediate outputs, and program outcomes. The elegance of the ToC paradigm is its explicitness: it forces a program designer to state each and every intermediary causal hypothesis of the program. In this way, having a ToC directs the development of the indicators to be measured: first, monitor the delivery of program elements; then, ensure that intermediary outputs have been achieved; finally, evaluate program outcomes. When well-executed, this type of monitoring allows designers to answer questions about why a program might not be performing optimally and provides them with the feedback needed for adaptive management.

While these elements of traditional ToCs and their accompanying indicators are critical, we propose that they are insufficient. This is because they implicitly assume critical psychological, social, and behavioral changes, without explicitly representing those changes within the ToC. However, the success of behavior change program development is entirely dependent on a proper understanding of the psychological and social environment in which an actor is behaving, and, more critically, on an understanding of how programming may change these environments or states. It is through changes to these states that we expect changes in behavior, which may, in turn, cause further changes to the social context. This means that quality behavior change programs based on that understanding of the psycho-social context include hypotheses about changes to the psychological and social characteristics of actors as well as their communities, including how interventions are expected to change these states, how these states affect behavior, and how that change in behavior cycles back to further change the social context. It is thus critical that the intermediate psychological and social indicators also be monitored. When measured properly, these indicators provide the same monitoring value as those in a traditional ToC, but also provide guidance for both diagnosing the cause of program success or failure, as well as facilitating adaptive management.

We employ the term *Psycho-Social Theory of Change* (PS-ToC) to describe the amalgamation of the above elements and more traditional ToCs. In a PS-ToC, between each element of an intervention and its intermediary behavioral outcomes lies the psychological or social changes assumed to have resulted from that intervention—i.e., the change that ultimately led to the behavioral outcome. It also allows for the representation of the dynamic relationship between behavior and psychological and social states, in which psychological and social states cause particular behaviors, which can, in turn, cause changes in those same psychological and social states. This explicit representation allows for a far greater understanding of why a program might be working or failing in terms of changes in the target psycho-social characteristics. It also allows for rapid adaptive management based on changes to those psycho-social factors long before an intermediate outcome might otherwise be observable. How this can be applied is discussed in detail in the *Frameworks for Indicator Monitoring and Adaptive Programming* section.

The application of a Psycho-Social Theory of Change requires the monitoring of three broad classes of indicators: behavioral, psychological, and social.

Behavioral Indicators

Monitoring changes in behavior is critical for any program claiming to target behavior change. In fact, for those behavior change programs where the link between the target behavior and environmental outcomes is already well documented, behavior may be evaluated as a final program outcome rather than an intermediate output measured in tandem with the environmental result of the intervention. Treating behavior as the program outcome is a particularly attractive option when a specific behavior is known to cause a longer-term environmental outcome, or when the measurement of behavior provides a less noisy signal of program success than the probabilistic impact a program may have on the target environmental outcome.

However, even when it may not be considered the final outcome, behavior is a critical intermediate output of any behavior change intervention. A proper PS-ToC, and therefore the indicators that follow from it, will not only include the final behavior that most closely causes the environmental outcome, it will also include each behavior identified as critical on the path to achieving that final behavioral goal. Once these behavioral indicators have been identified, they can be measured in a variety of ways (including direct observation, proxies, and self-reports), each offering unique benefits and costs. Please refer to the *Indicator Measurement* section for further discussion.

Psychological Indicators

Many program intervention elements aim to change behavior through the beliefs and preferences of their target actors. In traditional ToCs, these changes are often implicit, directly linking the intervention elements to the behavioral output. However, a PS-ToC makes explicit the intermediary step of belief and preference shift. By including this step, the PS-ToC acknowledges that even if the intervention element is delivered as directed, it may fail to change the necessary beliefs or preferences, and, therefore, fail to change behavior. The inclusion of psychological indicators tied to target beliefs and preferences is thus critical for understanding the degree to which a program was successfully implemented, and to identify adaptive management decisions that can or could be made.

Psychological Indicators in Farmer Training

An extension worker-led farmer training attempts to change three key target beliefs: (1) that overwatering is decreasing yields, (2) that watering three days a week is sufficient, and (3) that others in the community are watering less. All three of these beliefs are hypothesized to be the reasons that this training may lead participating farmers to waterless. They are, therefore, part of the PS-ToC, and psychological indicators to be incorporated into the monitoring and evaluation of that intervention component.

Social Indicators

Social indicators can be relevant at three distinct parts of a program's theory of change. First, social indicators can be key outcomes of a behavior change program. These can include what are commonly called objective outcomes, like how by restoring coastal habitats a program may aim to increase the livelihoods of those who rely on fisheries, as well as subjective outcomes, such as the program's effect on improving fishing communities' well-being. Second, a program may see changing social structure as both causing and being caused by behavior change—things like increasing trust in a particular institution or modifying the structure of a social network. Since changes

in both of these social indicator categories are critical for establishing why a program works and whether it has achieved its intended outcomes, it is critical for these social elements to also be captured in a PS-ToC and to serve as useful indicators for monitoring and evaluation.

Finally, social structures may represent the context in which an intervention takes place, even if this structure, in of itself, is not changed by said intervention. In other words, a particular social context may be critical for an intervention to work, but it might not itself be directly intervened on. This category is different from the two above, as those were expected to change as a result of the intervention. It can be important to measure these social contextual factors, as they can provide critical predictive information about whether an intervention is expected to work in a given context. For this reason, they should be represented in a PS-ToC. However, because they are not hypothesized to change over time, they should not be included as indicators for measuring program effectiveness or success.

Social Indicators in Farmer Interaction

A multi-community gathering of innovators is designed to change the social network farmers rely on for information on new environmentally-friendly farming techniques. The structure of target farmers' social networks, in relation to whom they rely on for information on farming techniques, is an integral part of the PS-ToC for the program. Changes to these social networks should be incorporated as an indicator of success relative said farmer-gathering program component. It is important to recognize that if the change to the social network is successful in changing behavior, that may result in a feedback loop further cementing that novel social network, a critical element to capture in the PS-ToC.

The same program may rely on trust in agricultural extension agents in order to convene the gathering. This is a key enabling social context that must exist for the program to succeed - it should be represented in the PS-ToC. However, because the program does not aim to change the degree to which the community trusts agricultural extension agents, it does not need to be monitored or evaluated for change.

Just as our understanding of behavioral and social science has advanced our understanding of the effective intervention design, so has it advanced our understanding of how we ought to monitor and evaluate them. Just as traditional ToCs forced program designers to be explicit about their previously implicit hypotheses, the PS-ToC forces that same explicitness with regard to changes in psychological and social factors. These explicit statements allow for the development of program-specific psychological and social indicators, which can then be incorporated into an understanding of the why of a program's success, as well as into rapid adaptive management efforts.

Figure 3 graphically depicts a PS-ToC example. It represents a component of the overall PS-ToC of Fish Forever, a global fisheries program that aims to increase fish stocks and community livelihoods in key coastal ecosystems for ten countries. The PS-ToC includes a mapping of various activities, as well as predicted changes to target actors' psychological and social states (these, in turn, are tied to key behaviors). By explicitly representing the intermediary social and psychological states a community is expected to experience, the PS-ToC provides clear guidance on the indicators that should be measured to assess whether a program is working, to diagnose how it might work better, and to create generalizable knowledge for future programming. The PS-ToC also includes explicit statements about the assumed enabling social conditions, which can be used in determining the intervention's applicability to a novel context.

It is important to recognize the bi-directional flow between behaviors and psycho-social states. While changes in psycho-social states cause changes in behavior, this process is dynamic, with those changes in behavior feeding back into the psychosocial context. This can be particularly pronounced and critical to capture in programs targeting social norm change, which is an inherently dynamic process whereby the changes in a segment’s behavior results in changes in the wider social context.

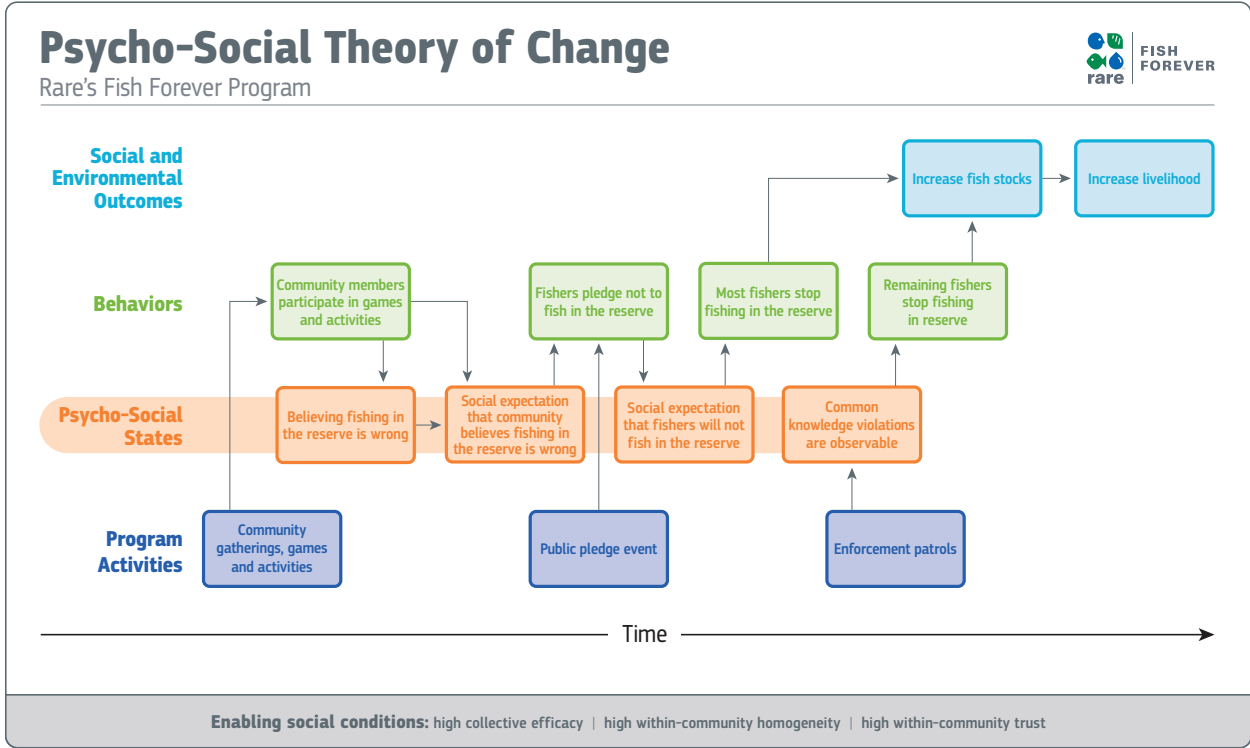


Figure 3: The psycho-social theory of change for Rare's Fish Forever program defines how program activities affect psycho-social states and behaviors, which in turn drive social and environmental outcomes. It also explicitly defines the enabling social conditions needed for the program to operate successfully.

Indicator Measurement

After establishing what indicators to measure, these need to be operationalized into assessable measures. Psychological indicators require introspective access, meaning we have to measure what is going on inside someone's head. This means that psychological indicators generally require self-report for measurement. This can be contrasted with behavioral indicators, which can be assessed with a variety of methods, including direct observation, where the behavior is directly measured; proxy measures, which assess outcomes which are assumed to be tightly related to the target behavior; and self-report measures, where the rate or intensity of a behavior is inferred through responses on instruments such as surveys. In this section, we review the benefits and drawbacks of each of these techniques and provide best practices to mitigate limitations.

Direct Observation

Direct observation represents the most straight-forward approach to measuring a behavioral indicator. However, direct observation is often limited by the observability of the behavior we want to monitor. Often, the behavior that program designers are interested in is the actual presence (or absence) of an illicit or private behavior. Such contexts require more creative, specialized design for successful direct observation.

As an example, take the use of toilets, a practice that has significant public health and environmental benefits. Toilet use has generally been seen as difficult—if not impossible—to measure through direct observation due to its inherently private nature. However, the introduction of PLUM devices, which use passive infrared motion detection to estimate toilet use rates, has provided a novel opportunity to directly evaluate toilet use behavior (Clasen et al., 2012). Used in this way, creative technological solutions can allow for the direct observation of practices previously seen as hard to assess.

Still, this does not mean that direct observation is without bias. Unless relying entirely on existing data sources collected without target actors' awareness, direct observation is subject to what is known as the Hawthorne Effect: where people's behavior is affected by the simple fact that they know they are being watched (Landsberger, 1958). Expanding on the previous example, the fact that PLUM devices allow for accurate observation does not address the problem that one might choose to behave differently depending on whether or not a PLUM device is installed on their toilet.

While direct observation is often seen as impossible for behaviors that are illicit, this is not always the case (given recent advances in remote sensing in particular; Kamminga et al., 2018). For example, the Eyes on the Seas initiative uses live satellite data to detect intrusion into marine protected areas in real-time (Kalinauckas, 2015). Technology has also made terrestrial poaching far more directly observable. For example, flying remote-operated craft over reserves in South Africa allows for the discrete and direct observation of rhinoceros poaching activity (Mulero-Pázmány et al., 2014). However, this style of direct observation is not always possible, and when it is, it may be prohibitively expensive. In these cases, proxy measures are reasonably employed to improve indicator measurement.

Proxy Measures

Proxy measures are those that, instead of directly measuring the behavior itself, measure some causal consequence. The prevalence of the behavior itself is then inferred from that consequence. These methods are most appropriate when direct observation is prohibitively costly or difficult to employ, and when the causal connection between the proxy and the indicator stands up to scrutiny.

An example of a successful proxy measure is assessing the amalgamation of various energy-saving behaviors through energy use, measured at the household meter (ex. Allcott & Rogers, 2014). A kilowatt-hour reduction is not

itself a behavior; instead, it is the consequence of a suite of behaviors that range from simple things like keeping the lights off all the way to replacing one's energy-inefficient appliances. Nevertheless, when an intervention targets these behaviors broadly, the measurement of household-level energy usage (relative to a valid control) represents a strong proxy measure for those behaviors—despite the measurement itself not being a direct measure of any of them. Many proxy measures have the added benefit that they are directly observable from already-collected data. Not only is this cheaper, it also means that any effect observed is unlikely to be the consequence of a Hawthorne style effect, where people are responding to having been unnaturally measured. Looking at an example from another domain, waste audits represent a similar tightly-linked behavior-proxy, making them particularly valid as proxy measures of consumption (Hoover, 2017). Waste audits generally involve examining the waste content from a household or institution, measuring the presence of different forms of waste. This measurement of waste (relative to a valid control) can be used to assess the relative effectiveness of waste reduction interventions despite never having to observe actual waste-reducing behaviors.

Some proxy measures are more dubious, however; often because of the role that other variables can play in influencing both the target behavior and any bias in proxy measurements. This is commonly observed in the use of enforcement logs, such as when arrest records are used to assess the prevalence of an illicit behavior. Imagine that an increase in enforcement incidents, like arrests, is observed. One cause may be a true increase in the prevalence of the illicit behavior. Another might be an increase in the presence of enforcement (which is, paradoxically, often the intervention itself). This means that one could observe an increase in enforcement actions because of an increase in the presence of enforcement, all while that same presence actually decreases the prevalence of the illicit behavior.

As demonstrated in these examples, proxy measures should be evaluated based on the degree to which they can be expected to result uniquely from the behavior of interest—and should be considered suspect if that is not the case. Additionally, the energy and waste examples, despite being strong cases, raise another common issue found in these measures: aggregation. While energy readings tell us that the household has adopted behaviors to reduce their energy consumption, they aggregate over a series of behaviors such that it is difficult to determine what unique actions were adopted among which members. Similarly, waste audits aggregate behavior across individuals and across time. This allows implementers to infer the average change in a behavior, but provides but a minimal understanding of how that change is distributed across individuals and situations.

Self-Report

Technically a form of proxy measurement, self-report measures ask the actor themselves to report on the prevalence of their behavior. These measures are subject to both random noise and bias due to cognitive and social factors. However, for many indicators, self-report may be the only practically available method. Additionally, many psychological and social indicators, by definition, require introspective access, meaning they must be assessed through self-report. Thankfully, a series of best practices can improve self-report measures by limiting, while not completely removing, the influence of these cognitive and social biases.

In Figure 4, Sudman et al. present a cognitive model of survey response, which helpfully illustrates the various opportunities for cognitive and social factors to add noise and bias to self-report measurement (1996).



Figure 4: Cognitive Model of Survey Response from Sudman et al. (1996)

The first four steps are particularly subject to cognitive bias, whereas response editing is primarily socially influenced. Designing valid self-report measures requires minimizing bias at each step in the process.

Example: Bias in Self-Report measures

We can explore each of these cases for bias introduction through the use of an example. Imagine we are attempting to measure the size of a fisher's average catch by asking them: "In a typical day, how many kilograms of fish do you normally catch?"

After being asked the question, the fisher must **interpret the language used**. For example, a fisher might interpret "a typical day" to mean only those days when the water is calm enough to fish, every day but Sunday, every day including weekends, etc. After interpreting the question, the fisher must then **retrieve the information from memory**. This is often done using cognitive shortcuts, which can bias their response. For example, a fisher might think back to a particularly salient day when they had a particularly large catch. They could alternatively think of yesterday's catch. Or perhaps vaguely think about how much a fisher would normally catch, rather than even thinking of their own one. To **translate that information into a judgment**, they must then conduct an internal calculation, perhaps by mentally counting the number of fish buckets they have on their boat and then estimating what proportion of those buckets was full on their last trip. They might also do this in a completely different way. And after that, they still have to **format their responses** to the response format that is made available to them. For example, if the question asks for an answer in kilograms, the respondent will need to estimate how many kilograms one bucket of fish represents and multiply that out. As has hopefully become clear, answering what might have looked like a simple question requires an extraordinary amount of mental effort—and every step risks further biasing one's answer. In order to improve the self-report process, items must, therefore, be made simple and unambiguous. While this advice may seem obvious on its face, meeting a sufficiently high degree of simplicity and clarity requires that each question be evaluated at each of the first four levels illustrated above: interpretation, retrieval, judgment, and format.

Unfortunately, simplicity and clarity are insufficient in dealing with the final question-answer step: **response editing**. Generally, response editing is the result of what we call the social desirability bias—that is, respondents are more likely to give answers that they believe are consistent with the surveyors' (and, more generally, society's) expectation about how they ought to have behaved. This bias is particularly pernicious given that it can be difficult to assess the magnitude of its effect. Various best practices and techniques do exist, however, that can minimize the bias' effect on participants' responses.

Minimizing social desirability bias starts long before the question is asked. The enumerator should ideally be from a neutral third party, rather than the implementing organization or the specific community of the respondent. While building rapport with the respondent, the enumerator should describe the goal of the data collection as being to better understand the participant, their beliefs, and motivations, rather than it being linked to an evaluation of the program. In the context of a particular self-report item, presenting both options as valid responses allows a respondent to feel less aberrant reporting their true behavior or beliefs. For example, one might add the preamble

“some people fish inside the MPA, and some people do not” before asking if the respondent does in order to validate both as reasonable answers. Additionally, answers can be made private, which minimizes the immediate uncomfortable nature of giving a socially undesirable response. This can be done in a variety of ways, such as having a respondent mark their answers on paper rather than verbally, or simply turning the enumerator’s tablet towards the respondent so that the enumerator cannot see it (yet allowing the respondent to give their report directly).

While these techniques may minimize bias through validation and privacy, they do not provide anonymity, which is often felt to be necessary for extremely socially sensitive questions (e.g., those relating to illicit behavior). Anonymity can be provided through anonymous polling. However, anonymous polling only allows inference at the aggregate level, both reducing statistical power and eliminating the ability to conduct individual-level analysis. Environmental social scientists have applied a number of advanced survey techniques, including randomized response and unmatched counts (Dalton et al., 1994; Warner, 1965), to assess the prevalence of illicit behavior all while preserving their ability to conduct individual-level inference (e.g., Bergseth et al., 2017; Solomon et al., 2007). Additionally, though offering incentives for truthful responses cannot eliminate the pull of social desirability, they can—at the very least—offset its effects. Such incentives could be utilized, for example, when the direct monitoring of all respondents’ behavior would be infeasible, but when occasional spot checks would otherwise be possible. In this case, an incentive can be tied to accurate reporting relative to the possible spot-check.

While self-report suffers from various possible threats to accuracy and precision, it may be the only possible option for logistical (such as cost) or logical (such as the indicator being psychological) reasons. In those cases, considering the various cognitive and social threats to validity outlined above, and designing to account for them, can improve the quality of self-report measures.

Framework for Indicator Monitoring and Adaptive Management

While monitoring and adaptive management remain frequent buzzwords for best practices, monitoring without clear intentions of how those data will be used for decision making undermines the overall effectiveness of an implementing organization. Similarly, some implementers collect monitoring data purely to satisfy the requirements of funding agencies, which represents a similarly ineffective—and unethical, as some have argued—approach (Gugerty & Karlan, 2018). However, the inclusion of psycho-social indicators derived from a program’s PS-ToC presents new opportunities for effective use of monitoring data in two ways. First, the data can be used for refined program diagnostics. Second, it can be used to inform live programmatic decisions that present an opportunity for net-reduction in overall program costs.

Psycho-Social Diagnostics

Traditional monitoring for adaptive management generally tracks the degree to which a program was delivered according to plan and may also include the tracking of intermediate behavioral outcomes. However, this tracking is insufficient for programmatic diagnostics and adaptive design. This is because it is missing the critical monitoring of the intermediate psycho-social indicators linking intervention components to behavioral outputs. In these traditional monitoring approaches, when the data indicate that the program elements were implemented according to plan, but the intermediate behavioral output was not achieved, this monitoring provides no insight into why the element failed to result in the intended behavioral outcome. It could be because the intervention element failed to change the targeted psychological or social state, or it could be because the change in that psychological or social state was insufficient for behavior change. This difference in diagnosis is critical for continued intervention development.

An alternative is the frequent pulse monitoring of the community on key psychological and social indicators. By combining these data with the traditional monitoring of program delivery and intermediate behavioral outcomes, a program designer can then assess where it is in the causal chain that breakdowns may be occurring (e.g., whether they are between the implementation and the psycho-social indicator, or between the psycho-social indicator and the intermediate behavioral output). This data can then be used to revise programming, either through the improvement of program elements to increase their ability to have the intended psycho-social effect, or through a shift in the theoretical understanding of the problem to focus on other psycho-social targets.

Live Programmatic Decision Making with Dynamic Programming

Monitoring data is most frequently used in a retrospective capacity to improve delivery in future instances. However, it is also used to ensure present programs are delivered as expected. This style of rapid management often focuses narrowly on the elements of the program that are directly delivered by the implementer, such as whether particular training sessions took place or particular tools were distributed. This corresponds to rapid response to program delivery indicators traditionally found in a program’s ToC.

The introduction of a PS-ToC presents an opportunity to conduct rapid response to psycho-social indicators to improve program efficacy and cost-effectiveness, a practice we refer to as dynamic programming. In a dynamic programming implementation, key psycho-social indicators of a program are assessed on a frequent basis throughout program delivery, a practice known as pulse monitoring. After an element of a program is implemented, the next round of pulse monitoring will indicate whether that program element has achieved the intended change in the psycho-social indicator. Pulse monitoring can also be used to assess feedback changes in the opposite direction in which changes in behavior cause changes in psychosocial states as hypothesized in the PS-ToC.

Tracking changes to the psycho-social landscape through pulse monitoring allows for rapid implementation decisions that can improve the effectiveness and cost of program delivery. While traditional methods of program delivery assume that for all units of intervention, whether they be individuals, communities, or geographies, the same dosage of an intervention is required to drive the needed change to psychological and social states. The dynamic programming approach acknowledges that there is heterogeneity in the dosage that an intervention requires.

This pulse tracking allows for two types of dynamic programming responses. If a unit is particularly responsive to one component of an intervention, the dosage of that component (for example, number of training sessions, the length of an incentive, etc.) can be reduced, as it has already achieved its desired effect. This has obvious cost-saving advantages. If a unit is particularly unresponsive to the standard dose of an intervention, either that dosage can be increased, or the program implementer can re-evaluate whether that particular component was actually appropriate for driving change in the target psychological or social state and whether expected feedback loops between behavior and psychosocial states are in fact present. While at first glance this may appear to be an increase in cost, but what would be even more costly is proceeding to the next step of an intervention without laying a proper foundation. Dynamic programming allows for the assurance that said foundation exists, thereby increasing program efficacy.

Methods for Pulse Monitoring of Psycho-Social Indicators

The pulse monitoring required for both improved psycho-social diagnostics and dynamic programming presents an additional cost to an implementing program. It is therefore important to identify methods that are scalable and cost-effective to be applied across programs, geographies, and social groups.

While many programs conduct in-person, pre-post interviews of program participants that can assess change in psycho-social indicators, these do not provide information useful for rapid programmatic responses. It would also be prohibitively costly for most programs to conduct frequent in-person interviews that gauge changes on these dimensions. We, therefore, present a scalable, technologically-assisted alternative.

Over the past decade, mobile phone ownership has skyrocketed across low-income countries. In 2019, Pew reported that the median emerging-economy country has almost 80% mobile phone penetration among adults. While remaining differential access to these phones based on a variety of social factors including income, gender, and social class needs to be accounted for, this high prevalence makes mobile phones an ideal candidate for scalable psycho-social indicator monitoring. This is in contrast to the prevalence of smartphones in emerging economies, which stands at only 45%. This discrepancy means that any measurement aiming for representative coverage needs to be administered over voice or text message, rather than through smartphone apps.

Automated text message surveying and interactive voice response (IVR) allow for the inexpensive large-scale monitoring of psycho-social indicators within a population. Researchers working in low-income contexts have found that both methods can be effective for high-frequency data collection, but that IVR tends to result in higher response rates as compared to text messages (Ballivian et al., 2015). By collecting these psycho-social measures on a regular basis over the life of the program, the implementer can track their changes live over time, and make informed rapid program delivery decisions.

The Development of a Dynamic Programming Framework for Coastal Fisheries Management

Rare's Fish Forever program described earlier is presently integrating pulse monitoring for dynamic programming into its 10-country program. As each site in the program presents unique programmatic needs, it presents an excellent opportunity to optimize differing dosages and delivery modes for the program's various components.

The Fish Forever community engagement program can be divided into three stages outlined in the Theory of Cooperative Behavior Adoption: generating collective demand, coordinating a shift in behavior, and strengthening the social norm (Thulin, 2020). This resulted in the following key psycho-social indicators in the Fish Forever PS-ToC for the key behavior of only fishing outside the reserve area.

Generating Collective Demand:

- Psychological: Believing that it is wrong to fish in the reserve
- Social: Common knowledge that the large majority of community members believe it is wrong to fish in the reserve

Coordinating a Shift in Behavior:

- Social: Common knowledge that all fishers in the community will no longer be fishing in the reserve
- Behavioral: Fishers are only fishing outside the reserve

Strengthening the Norm

- Psychological: Believing that if one were to fish in the reserve, the rest of the community would find out
- Social: Common knowledge that all members are expected to stay outside the reserve, and that intrusion would be socially sanctioned

It is critical for the PS-ToC logic of the Fish Forever program that these three steps occur in sequence. It is therefore important that the psychological and social indicator thresholds be met in the '*Generating Collective Demand*' step before proceeding to the '*Coordinating a Shift in Behavior*' step. Given the incredible diversity of sites in which the Fish Forever program is deployed, implementers cannot rely on the same timeline for progress on these steps in every site.

To address this challenge, Rare partnered with the telecommunications service company EngageSpark to develop an IVR based solution that would measure each of these six indicators on a bi-weekly basis among a panel of respondents within each community in the Fish Forever program. This bi-weekly feedback is designed to directly feed into the program delivery decisions of the implementer, telling them when it is the appropriate time to proceed to the next step in the cooperative behavior adoption process. When a community does not experience the expected changes in target indicators, a standardized qualitative methodology is employed to determine why the programming is not working. This may result in either a change in the method of delivery,

or simply, in additional time spent on those components of the intervention specific to that community until targets are achieved.

The Fish Forever dynamic programming platform is currently undergoing testing, with the goal of program-wide rollout in 2021.

Figure 5 graphically depicts how pulse monitoring can be incorporated into live program delivery decisions, using Rare's Fish Forever program as an illustrative example. Throughout each phase of the program, the pulse monitoring assesses the degree to which target psychosocial states have been achieved. If they have not, and the site is therefore behind schedule, the implementation team then conducts a rapid qualitative assessment to determine why the intended psychosocial states have not been achieved. This assessment is then incorporated into localized revisions to the program activities, which are then executed until the target state is reached.

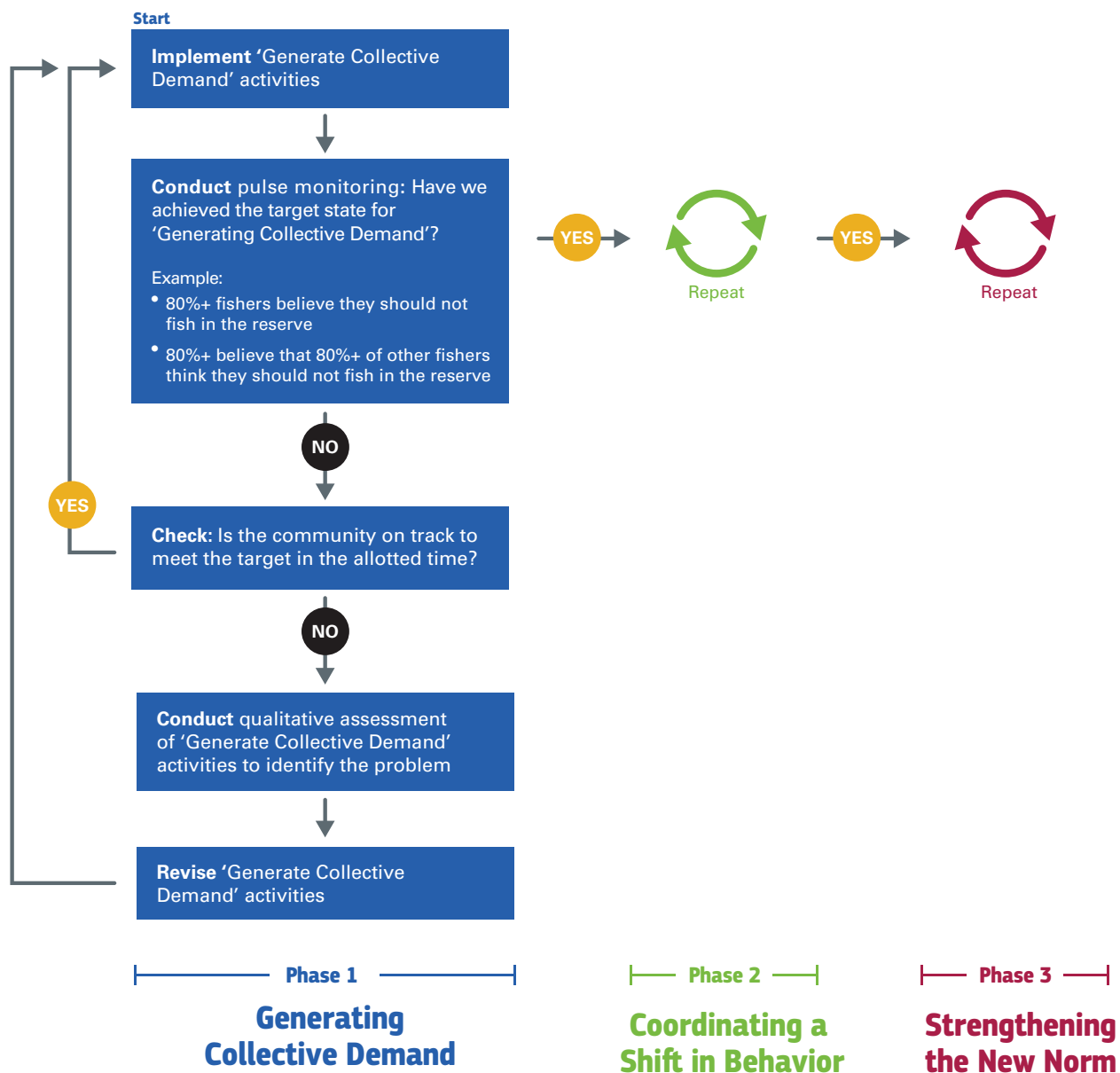


Figure 5: Flowchart representing the steps in the Dynamic Programming process of Rare's Fish Forever program. Dynamic Programming represents the inclusion of pulse monitoring, where the psychological and social state of the community is continually tracked over time to improve live programmatic decision making such as when to advance to the next phase in a program.

Frameworks for Evaluating Changes in Indicators

Evaluation serves a critical role in a behavioral intervention lifecycle. After the development and testing of a program, rigorous evaluation allows for strong causal claims about the relative effectiveness of the intervention, as compared to either a control or a common “standard of care” intervention. This is an important step for a number of reasons. First, without robust estimates of the treatment effect of an intervention, it is impossible to calculate the cost-effectiveness of a program. Second, it is critical from an ethical perspective: properly determining that a program is causing an observed change is a necessary input into whether the scaling up of that program is ethical in terms of both the deployment of resources as well as the way we treat target communities.

It is important to acknowledge that this is not to say that every deployment of every intervention needs to be evaluated against a counterfactual. Similar to how many medicines have well enough documented effects that we no longer need to assess their effectiveness, some behavior change programs may be similar. The key question to answer is whether a particular program has been demonstrated in a sufficiently similar context to confidently estimate the size of its effect without any evaluation. The most common situation in which this is the case is when a program has been demonstrated as successful for a random subset of a target population, and then, is deployed for that entire population. It is almost never the case when a program is applied to a new social or ecological context that it will be as successful as before.

There are various designs for causally inferring the treatment effect of an intervention. Each of these methods is fundamentally attempting to ask the question: how would those in the population have responded to a particular intervention compared to a counterfactual world where no treatment was given at all? This difference tells designers what the causal effect of that intervention is. In this section, we aim to aid in the initial evaluation of these options for a given program by giving an overview of their logic as well as their relative costs and benefits to other possible evaluation designs.

Randomization Based Methods

Randomization is commonly seen as the gold standard in causal inference evaluation as it relies on the fewest assumptions for unbiased estimates. We therefore present randomization-based methods first, as all other methods of causal inference will be compared to them in terms of what additional assumptions they require.

In this section we review Randomized Controlled Trials (RCTs) and Clustered Randomized Trials (CRTs). RCTs are normally used when an intervention can be randomized at the individual or household level. CRTs are used when interventions can only be randomized at some group level, such as with a community or region, but the effect of that intervention can still be measured at the individual or household level.

Randomized Controlled Trials

Fundamentally, RCTs require units, often individuals, to be randomly assigned to a particular treatment condition. These can be the intervention of interest, a standard control condition (meaning no treatment), or some other treatment to which the treatment of interest is compared (this is sometimes referred to as an active control). To evaluate the effectiveness of an intervention, the average value of the control group’s indicator of interest is subtracted from the average value on the indicator in the treatment group. The remaining value is the average treatment effect, representing the change in the indicator that can be causally attributed to the treatment.

The strength of RCTs rests on a fundamental assumption of randomization. At its core, this assumption is that every unit in the population of interest has an equal chance of being selected to participate in the study, and every unit in the population that has been selected has an equal chance of being assigned to each treatment condition. This means that if a partner organization only wishes to work in particular communities, then any RCT run within that setting is only technically generalizable to those particular communities rather than to the wider population. It also means that if any preference is given for any units to be assigned to the treatment versus the control condition, the validity of inferences is undermined.

Method for Cost Savings in RCTs: Unequal Group Sizes

Those who are designing evaluation protocols often assume that the most cost-efficient way to do so is to allocate the same number of units to the treatment and control conditions. However, this is not always the case. Specifically, this does not hold true when the cost of recruiting and evaluating an intervention unit is different from the cost of a control unit. For example, an intervention evaluating the effectiveness of delivering clean cookstoves may find that the cost of delivering, training, and surveying in the treatment condition is \$90 per participant, whereas the cost of only surveying in the control condition is \$10. In this case, the most cost-efficient allocation is to recruit three times more respondents in the control condition than in the treatment condition. The method for calculating this optimal allocation can be found in List et al. (2011).

Clustered Randomized Trials (CRTs)

Discussion of RCTs, like the above, generally focuses on their application to programs that target individuals, and critically, that can be randomized at that individual level. This is highly applicable for those behavior change programs that are suitable for individual or household level randomization. However, due to their fundamentally geographic and often community-focused nature, many environmental programs are not suitable to be delivered at that level of granularity, and instead, must be administered at the community or regional level. CRTs present a randomized design framework suitable for assessing such interventions.

The setup of a CRT is quite similar to an RCT: clusters, such as communities, are randomly assigned to either receive the intervention of interest or some comparison control. The key difference with RCTs rests on randomization at the cluster rather than individual level. And this has significant implications for the study design: depending on the degree of similarity between individuals within each cluster, CRTs can require a much large number of clusters. Simply having two communities with individuals from one community assigned to one treatment and individuals in one other community assigned to another treatment, even with a high number of individuals monitored within each community, will never be sufficient for valid statistical inference.

Thankfully, if a sufficient number of clusters are recruited to participate, CRTs present causal inferences that are as strong as those of RCTs—at least for estimating the average treatment effect of an intervention given the minimal assumptions tied to randomization.

Quasi-Experimental Methods

While randomization-based experimental methods offer strong causal inference with minimal assumptions, randomization is not always an available evaluation strategy. This has led to the development of quasi-experimental

strategies that attempt to infer the causal effect of an intervention without randomization, albeit with additional assumptions that are generally impossible to assess directly.

Difference-in-Difference

Difference-in-difference is a commonly employed strategy when the assignment to a condition cannot be randomized, but when baseline data is available for both those who receive the treatment as well as for a group that was not assigned to receive the treatment but that appears otherwise similar. Instead of simply comparing outcome indicators after the treatment has gone into effect, difference-in-difference compares the change in indicator measures of the treatment group with the changes in that of the control group. The average treatment effect of the intervention is then estimated as the change in an indicator of the control condition value subtracted from the change in the indicator's value in the intervention condition.

This comparison of differences, rather than just raw outcomes, is critical for the logic of difference-in-difference. Because of this focus on change, difference-in-difference does not need to assume that groups assigned to the intervention and those assigned to the control are at a similar level on the outcome indicator. Instead, the most critical assumption of difference-in-difference is parallel trends. This assumes that, in the absence of any intervention, those in the intervention group would experience the same amount of change in outcome indicators as those in the control group. This assumption is generally not testable, however, and must be argued on a case by case basis.

Matching

Matching techniques are an additional method for overcoming the hurdle of not being able to randomize units to a control and intervention condition. Instead of randomizing, matching attempts to build a control condition by identifying units that are similar to each of the treatment units—doing so relative to a set of observable characteristics.

There are multiple methods for finding matches. Two common methods are exact matching and propensity-score matching (PSM). Exact matching relies on being able to identify units to use as the control group which show the exact same values on each of the observable characteristics (such as someone's gender, SES, distance from a resource) for the treatment units, except for the fact that they did not get the intervention. Alternatively, PSM focuses on those characteristics that are likely to predict receiving the intervention—matching people based on their likelihood of receiving the intervention rather than on the underlying characteristics that define it. This means that for each unit evaluated in the treatment condition, there is a unit in the control condition that, because of its baseline characteristics, would have been expected to be in the treatment condition (i.e., had a similar likelihood) but was not.

The drawback of matching methods is that they are only matched on observable characteristics, and these almost never include all of the characteristics that could explain why a unit received the treatment instead of being assigned to the control. This failure to include all (not just observable) explanatory variables means that any matched comparison could be subject to bias because of those unobservable characteristics yet unaccounted for.

Pre-post Comparison

Pre-post comparison involves estimating the effect of an intervention by simply subtracting the average baseline value of the outcome indicator from the average post-intervention value of the outcome indicator. While this technique is commonly used to evaluate environmental interventions, it rarely ever renders valid estimates.

Pre-post comparison makes the incredibly strong assumption that no other factor influenced the value of the outcome indicator between the pre and post measurements except the intervention itself. This assumption is rarely ever true, and even more rarely ever justifiable.

One case in which this may hold for a behavioral intervention is when a new practice is being introduced, and no other actors are operating in the area which might introduce that practice. This could be seen, for example, in the adoption of a novel product, where that product is available from those implementing the intervention and no other sources. However, this narrow case rarely represents the practices that environmental behavioral interventions seek to address.

Additional Methods

While the randomized and quasi-experimental methods described above represent the large majority of techniques used to evaluate environmental programs, they are not exhaustive. Additional econometric techniques, which are often used to evaluate policy interventions, are also sometimes applicable, including regression discontinuity, instrumental variable analysis, and synthetic controls. For those wishing to dive further into these techniques, Cunningham (2021) provides an excellent introduction.

Disaggregating Impact Across Social Differences

Due to behavior change programs targeting specific barriers and motivations, and people with different social identities having different barriers and motivations, behavior change programs can have dramatically different effects for different social groups. Traditional methods of evaluation simply aggregate the entire sample across these groups together, returning the average treatment effect of an intervention. However, this average hides the nuance of whom is being affected by an intervention. It is therefore critical to analyze and report disaggregated estimates of program effects, which show how an intervention may have impacted different social groups differently. A key factor in identifying on what dimensions an effect should be disaggregated is to return to the program's foundational analysis of the socio-ecological context to identify any groups which might experience different motivations and barriers due to their social position. This disaggregated analysis is critical for two reasons. First, it allows for far more nuanced learning to incorporate into adaptive management of a program to increase effectiveness. Second, it allows implementers to evaluate the equitability of program outcomes.

Cross-Context Generalizability

The methods above provide varying levels of strong causal inference. However, all of these methods are subject to the constraint that the estimated intervention effects they identify are only representative of the population from which they were drawn. This issue is known as the generalizability puzzle: if something works in one socio-ecological context, will it apply in another? Thankfully, recent advancement in this space helps us begin to answer that question. Too often, those considering the scaling of a program focus exclusively on geography, asking whether a particular intervention was tested in the same country or region. However, in order to answer the question of whether an intervention will work in a particular context, instead of asking where an intervention has worked, it is critical to ask why the intervention worked (Bates & Glennerster, 2017).

This returns to the earlier discussed Psycho-Social Theory of Change. While a traditional ToC gives very little insight into why an intervention worked, a PS-ToC explicitly states the why as a hypothesis. By doing so, one can then ask the question, does this why apply to this new context?

Bates and Glennerster lay out four steps for evaluating the generalizability of an intervention to a new context, adapted here:

1. Does the intervention have a Psycho-Social Theory of Change?
2. Given the PS-ToC, do the enabling conditions, including socio-ecological conditions, hold?
3. Given the PS-ToC, are you confident that the underlying psycho-social logic linking intervention components to behavioral outputs applies to the members of the new target population?
4. What is the evidence that the implementation process can be carried out with high fidelity?

While traditional evaluation will not tell designers whether an intervention will generalize, this four-step process integrates evaluation with our mechanistic understanding of the social and psychological context to allow for a more generalized understanding.

Conclusion: Recommendations on Indicators For Behavior Change Programming

We conclude with a synthesis of the content above into a series of best practices for the development, monitoring, and evaluation of indicators for behavior change programming.

1. In addition to the traditional components of a theory of change, the development of behavior change programming should include the adoption of a psycho-social theory of change (PS-ToC). The PS-ToC should include explicit representation of the psychological and social changes expected as the result of each element of a program between that element of the program and the behavioral output, as well as the necessary enabling conditions.
2. Program indicators for behavior change programming should include indicators for psychological and social elements of the PS-ToC linking intervention components to behavioral outputs.
3. Behavioral indicators are ideally measured through direct, unobtrusive observation. However, when direct observation is infeasible, carefully selected proxy measures of behavior and carefully drafted self-report measures may also be valid. Proxy measures should be closely, logically linked to the target behavior, and survey measures should be administered and drafted in a manner that accounts for best practice in addressing cognitive and social response biases.
4. Measures of psychological and social indicators often require actors to reflect on their own mental states and subjective view of their social circumstances that influence their behavior, and therefore can often only be measured through survey items. These, too, should be drafted applying best practices to reduce cognitive and socially biased responses.
5. Behavior change programming should adopt rapid, pulse-style, assessment of changes in psychological and social states of the communities in which they are deployed. This will allow for better adaptive management that improves future behavior change programming, as well as allowing for dynamic programming where an intervention can be designed to rapidly incorporate live psychological and social state information to improve a program's efficacy and cost-effectiveness.
6. Randomized evaluation, either through randomized controlled trials or clustered randomized trials, is the method of estimating program effectiveness that requires the fewest statistical assumptions, and should, therefore, be adopted for the evaluation of behavior change programs whenever possible. Quasi-experimental methods, such as difference-in-difference and matching, are also valid methods for inference but require additional, often untestable assumptions. Pre-post comparisons are rarely logically justified, and are very likely to yield statistically biased estimates. They should thus be avoided in the vast majority of cases.
7. Generalizability of a program across contexts cannot simply be derived from the statistical evaluation of a program. Nor can we assume that because a program has worked in one context, it will generalize to neighboring geographies. Instead, we must rely on our understanding of the enabling social and material environment as well as the program-specific psychological and social linkages between program intervention and behavioral output, which describes the why a program works. In order to evaluate whether a program will generalize to a particular novel context, it must be believed to be sufficiently similar for that same why to apply.

References

- Allcott, H., & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, 104(10), 3003–3037. <https://doi.org/10.1257/aer.104.10.3003>
- Ballivian, A., Pedro Azevedo, J., & Durbin, W. (2015). Using Mobile Phones for High-Frequency Data Collection. In *Mobile Research Methods: Opportunities and challenges of mobile research methodologies*. Ubiquity Press. <https://doi.org/10.5334/bar.c>
- Bates, M. A., & Glennerster, R. (2017). The Generalizability Puzzle. *Stanford Social Innovation Review*. https://ssir.org/articles/entry/the_generalizability_puzzle
- Bergseth, B. J., Williamson, D. H., Russ, G. R., Sutton, S. G., & Cinner, J. E. (2017). A social–ecological approach to assessing and managing poaching by recreational fishers. *Frontiers in Ecology and the Environment*, 15(2), 67–73. <https://doi.org/10.1002/fee.1457>
- Clasen, T., Fabini, D., Boisson, S., Taneja, J., Song, J., Aichinger, E., Bui, A., Dadashi, S., Schmidt, W.-P., Burt, Z., & Nelson, K. L. (2012). Making Sanitation Count: Developing and Testing a Device for Assessing Latrine Use in Low-Income Settings. *Environmental Science & Technology*, 46(6), 3295–3303. <https://doi.org/10.1021/es2036702>
- Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2010). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009: *American Journal of Evaluation*. <https://doi.org/10.1177/1098214010389321>
- Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press.
- Dalton, D. R., Wimbush, J. C., & Daily, C. M. (1994). Using the Unmatched Count Technique (uct) to Estimate Base Rates for Sensitive Behavior. *Personnel Psychology*, 47(4), 817–829. <https://doi.org/10.1111/j.1744-6570.1994.tb01578.x>
- Gugerty, M. K., & Karlan, D. (2018). *The Goldilocks Challenge: Right-Fit Evidence for the Social Sector* (1 edition). Oxford University Press.
- Hoover, D. (2017). Estimating Quantities and Types of Food Waste at the City Level. *Natural Resources Defense Council*.
- Kalinauckas, A. (2015). Eyes on the seas. *Engineering Technology*, 10(3), 68–69. <https://doi.org/10.1049/et.2015.0353>
- Kamminga, J., Ayele, E., Meratnia, N., & Havinga, P. (2018). Poaching Detection Technologies—A Survey. *Sensors*, 18(5), 1474. <https://doi.org/10.3390/s18051474>
- Landsberger, H. A. (1958). Hawthorne Revisited: Management and the Worker, Its Critics, and Developments in Human Relations in Industry.
- List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, 14(4), 439. <https://doi.org/10.1007/s10683->

- Mulero-Pázmány, M., Stolper, R., Essen, L. D. van, Negro, J. J., & Sassen, T. (2014). Remotely Piloted Aircraft Systems as a Rhinoceros Anti-Poaching Tool in Africa. *PLOS ONE*, 9(1), e83873. <https://doi.org/10.1371/journal.pone.0083873>
- Solomon, J., Jacobson, S. K., Wald, K. D., & Gavin, M. (2007). Estimating Illegal Resource Use at a Ugandan Park with the Randomized Response Technique. *Human Dimensions of Wildlife*, 12(2), 75–88. <https://doi.org/10.1080/10871200701195365>
- Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). Thinking about answers: The application of cognitive processes to survey methodology (pp. xiv, 304). Jossey-Bass.
- Thulin, E. (2020). Cooperative Behavior Adoption Guide: Applying Behavior-Centered Design to Solve Cooperative Challenges. Rare.
- Warner, S. L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 60(309), 63–69. <https://doi.org/10.1080/01621459.1965.10480775>



Rare inspires change so people and nature thrive. Conservation ultimately comes down to people – their behaviors toward nature, their beliefs about its value, and their ability to protect it without sacrificing basic life needs. And so, conservationists must become as skilled in social change as in science; as committed to community-based solutions as national and international policymaking.

The Center for Behavior & the Environment at Rare is translating science into practice and leveraging the best behavioral insights and design thinking approaches to tackle some of the most challenging environmental issues. Through partnerships with leading academic and research institutions, they are bringing the research into the field to connect the next generation of behavioral scientists with practitioners on the front lines of our greatest environmental challenges.

To learn more, visit behavior.rare.org



The Global Environment Facility (GEF) was established on the eve of the 1992 Rio Earth Summit to help tackle our planet's most pressing environmental problems. Since then, the GEF has provided close to \$20.5 billion in grants and mobilized an additional \$112 billion in co-financing for more than 4,800 projects in 170 countries. Through its Small Grants Programme, the GEF has provided support to nearly 24,000 civil society and community initiatives in 133 countries.

The Scientific and Technical Advisory Panel (STAP) comprises seven expert advisers supported by a Secretariat, which are together responsible for connecting the GEF to the most up to date, authoritative, and globally representative science. The STAP Chair reports to every GEF Council meeting, briefing Council members on the Panel's work and emerging scientific and technical issues.