Gates Open Research

Gates Open Research 2019, 3:886 Last updated: 06 JAN 2020

METHOD ARTICLE

CUBES: A practical toolkit to measure enablers and barriers to behavior for effective intervention design [version 1; peer review: 2 approved, 1 approved with reservations]

Elisabeth Engl [1], Sema K. Sgaier [1-3]

[1]Surgo Foundation, Washington, District of Columbia, 20001, USA
[2]Department of Global Health & Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[3]Department of Global Health, University of Washington, Seattle, Washington, USA

v1 **First published:** 18 Mar 2019, **3**:886 (https://doi.org/10.12688/gatesopenres.12923.1)

Latest published: 06 Jan 2020, **3**:886 (https://doi.org/10.12688/gatesopenres.12923.2)

Abstract

A pressing goal in global development and other sectors is often to understand what drives people's behaviors, and how to influence them. Yet designing behavior change interventions is often an unsystematic process, hobbled by insufficient understanding of contextual and perceptual behavioral drivers and a narrow focus on limited research methods to assess them. We propose a toolkit (CUBES) of two solutions to help programs arrive at more effective interventions. First, we introduce a novel framework of behavior, which is a practical tool for programs to structure potential drivers and match corresponding interventions. This evidence-based framework was developed through extensive cross-sectoral literature research and refined through application in large-scale global development programs. Second, we propose a set of descriptive, experimental, and simulation approaches that can enhance and expand the methods commonly used in global development. Since not all methods are equally suited to capture the different types of drivers of behavior, we present a decision aid for method selection. We recommend that existing commonly used methods, such as observations and surveys, use CUBES as a scaffold and incorporate validated measures of specific types of drivers in order to comprehensively test all the potential components of a target behavior. We also recommend under-used methods from sectors such as market research, experimental psychology, and decision science, which programs can use to extend their toolkit and test the importance and impact of key enablers and barriers. The CUBES toolkit enables programs across sectors to streamline the process of conceptualizing, designing, and optimizing interventions, and ultimately to change behaviors and achieve targeted outcomes.

Keywords

Intervention design, implementation science, behavior change, behavioral drivers, behavioral models, research methods, global health, global development.

Open Peer Review

Reviewer Status ✔ ? ✔

	Invited Reviewers		
	1	**2**	**3**
REVISED **version 2** published 06 Jan 2020			
version 1 published 18 Mar 2019	✔ report	? report	✔ report

1 **Hope Hempstone**, United States Agency for International Development (USAID), Washington, USA
Angela Brasington, United States Agency for International Development (USAID), Washington, USA

2 **Rashida A. Ferrand**, London School of Hygiene and Tropical Medicine, London, UK
Biomedical Research and Training Institute, Harare, Zimbabwe

3 **Neela A. Saldanha**, Ashoka University, Sonipat, India

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Sema K. Sgaier (semasgaier@surgofoundation.org)

Author roles: Engl E: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Sgaier SK**: Conceptualization, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was funded by the Bill and Melinda Gates Foundation and the Surgo Foundation.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Engl E and Sgaier SK. **CUBES: A practical toolkit to measure enablers and barriers to behavior for effective intervention design [version 1; peer review: 2 approved, 1 approved with reservations]** Gates Open Research 2019, **3**:886 (https://doi.org/10.12688/gatesopenres.12923.1)

First published: 18 Mar 2019, **3**:886 (https://doi.org/10.12688/gatesopenres.12923.1)

Introduction

Interventions that aim to shift what people do, and the choices they make, are a major focus of global programs and policy. With aggressive targets set by the United Nation's Sustainable Development Goals and limited funding for global development, programs must leverage limited resources effectively and efficiently[i]. To achieve global development outcomes, such as reducing maternal mortality or the number of HIV infections, a plethora of diverse interventions have been designed and implemented across the world that ultimately all aim to shift behavior. Examples include increasing the demand for polio vaccinations in India through mobilization activities[1], developing checklists for nurses in Uttar Pradesh, India, to increase adherence to labor and delivery guidelines[2], and creating incentives to drive voluntary male circumcision for HIV prevention in Kenya[3].

Not all these examples have resulted in successful and sustainable behavior change at the levels needed to have the desired impact on development outcomes. Creating lasting change through successful interventions is hard. First, it requires a thorough understanding of *why* a target behavior is currently not occurring in a given context. Designing and evaluating interventions in the field without a thorough understanding of the underlying drivers of the target behavior in question can be an inefficient use of time and resources[4]. Second, since people's choices and behaviors are influenced by many factors, a single intervention is often insufficient to drive change. For instance, a failure of malaria net uptake may be rooted in a lack of availability or accessibility. But this external context is just one part of the picture: beliefs on the part of the end-user, for example about the benefits and risks of the nets, along with many other factors, may also influence decision-making. All these factors must be addressed through a well thought-through set of complementary interventions (an intervention portfolio) to significantly improve the usage of bed nets. Third, not all people are the same. Sub-groups of people within the target population can be differentiated by the varying drivers behind their behavior, necessitating a different set of interventions to be targeted at each subgroup[5,6].

The right levers to focus on, and the portfolio of targeted interventions to scale, are often far from obvious. Programs need a holistic and practical behavioral framework that accounts for and structures all types of barriers and enablers of behavior. Many models of behavioral enablers and barriers exist, but most focus either on systemic drivers, without addressing how individuals can be motivated to respond to such changes[7–9], or on people's beliefs, personality characteristics or cognitive biases, neglecting context[10–19]. Several approaches, such as the COM-B behavior change wheel, the Fogg model, and MINDSPACE, focus most strongly on the appropriate types of interventions to change behavior[20–22]. However, no current behavior change framework helps programs select the right research tool to assess an intervention's components. Programs need a repertoire of validated methods that will help them assess distinct enablers and barriers of behavior change in the field, because not all methods are effective at measuring each type of enabler and barrier, and only a limited set of methods are used in the development sector today.

In this paper, we provide programs with a practical two-part toolkit to help programs design an effective portfolio of interventions. We call the toolkit CUBES: to Change behavior, Understand Barriers, Enablers, and Stages of change. First, we present an evidence-based framework for understanding behavior. The framework synthesizes stages of change, contextual and perceptual drivers (which can act as enablers or barriers), and layers of influencers, using evidence from multiple sectors. The components were first articulated and applied in the voluntary medical male circumcision program[23], and we later refined the framework through a thorough evaluation of existing behavioral models, and by testing its applicability and practicality in several large-scale development programs. Second, to help programs generate actionable insights into the components of the framework in their own context, we recommend a set of research methods from various sectors and detail their strengths and weaknesses, expanding the methodological toolkit of qualitative interviews and quantitative surveys that many programs use by default. We support this with a decision tree to aid the choice of research method according to the practitioners' specific development program and context.

Real-world programs aim to drive change in complex dynamic systems of people, places, and information channels, and CUBES can be applied at any level in the system. To illustrate the usability of the proposed approaches, we present a case study showing how the CUBES framework and the methods toolkit were applied in a large-scale program for voluntary medical male circumcision in Africa. While we have developed this approach through the lens of our programs in global development, its principles can be applied to any behavior change context.

Methods

Constructing a best-practice framework of behavior

Grounding intervention design in a comprehensive and actionable behavioral framework is important. Many such models exist, with varying levels of evidence supporting their components. We surveyed models that were a) most influential and b) had an evidence base confirming that they predict behavior, rather than a comprehensive survey of all models in existence. We defined 'influential' as having been used to guide behavioral-intervention design across sectors, including health behaviors, and 'evidence-based' as the existence of original research evaluating the predictive power of individual components of a model (such as 'perceived severity of risk'[24] or 'conscientiousness'[12]) on behavior.

We began with a list of models known to the authors, then surveyed key approaches cited in these models, then searched PubMed and Google Scholar for the following terms: 'behavior' AND (model OR framework OR drivers OR barriers OR facilitators OR enablers), and focused on the first five pages of

[i]United Nations. About the sustainable development goals. Available from: https://www.un.org/sustainabledevelopment/sustainable-development-goals/.

results in the search engines. The main search was performed between October and December 2016, with additional targeted searches until December 2018. We identified 17 models fitting the criteria, each focusing on a different set of behavioral drivers (Table 1). The drivers for which we found experimental evidence of moderate to high predictiveness on behavior were then placed into a framework. We subjected the drivers included in the framework to critical review by four experts in behavioral science, health psychology, and the development sector, focusing on the drivers' comprehensiveness and applicability to global development. Finally, we applied the framework to design research in our own large-scale programs (see sample case study in this paper) to test for actionability and to further refine its components.

Assessing and developing a curated set of research methods to measure drivers

To help programs select appropriate research tools to capture enablers and barriers to behavior, we surveyed methods used across disciplines, using literature research and expert conversations. Unless otherwise mentioned below, we then applied and tested these methods in various combinations in our own large-scale development programs to assess each method's feasibility, strengths and weaknesses in insight generation. In Zambia and Zimbabwe, we investigated voluntary medical male circumcision[4,5,23], and in different areas of India we conducted programs on household behaviors relating to maternal and child health (family planning, antenatal care, institutional delivery, postnatal care), tuberculosis care-seeking behaviors, and health-care provider and front-line worker behaviors within medical facilities and communities (unpublished reports).

Results

Cubes: a practical framework of behavior

Following our review of influential models of behavior, we distilled their most evidence-based components into a practical behavioral framework that programs can use to evaluate existing evidence, conduct research to close evidence gaps, and ultimately design interventions to match barriers to behavior (Figure 1). The CUBES framework articulates three critical components of behavior change. First, the path toward a target behavior consists of a series of distinct stages. Second, the progression through each of these stages is influenced by a set of contextual and perceptual drivers. Third, these barriers and enablers may be transmitted to the individual through influencers (such as friends, family, or community members), either directly or through media channels. Below we outline each component and the contributions of the behavioral models surveyed (Figure 1).

Stages of the behavioral change process

Both contextual and perceptual drivers influence whether individuals possess the knowledge needed for behavior change, intend to act, or are already acting. These drivers either hinder or facilitate progression along the stages of change.

In global development, many intervention programs focus on enhancing awareness (passive knowledge) and skills (active knowledge). However, it is possible to be aware of an option, and even have the skills to do something, without intending to take advantage of it. For example, in Uttar Pradesh, India, increasing nurses' skills did not always increase their practices accordingly[25]. Clearly, knowledge does not equal action. Therefore, understanding where people are on the pathway to behavior, and why they are not moving forward, is key to designing interventions that can move people toward action. This is an essential first step for program designers to orient themselves when evaluating behavior.

The Transtheoretical Model[13] proposed a series of relatively rigid stages: precontemplation (being unaware), contemplation, preparation, action, maintenance, and termination, and adds a set of interventions ('processes of change') to help an individual progress from one stage to the next. However, the efficacy of assigning individuals to very detailed stages has been called into question in several systematic reviews[26–28]. In line with the Health Action Process Approach[16], the CUBES framework therefore divides the behavioral change process more simply into three stages of knowledge (ending with the necessary awareness or skills to engage in a behavior), intention, and action. While the path to action and beyond can be understood as a sequence, people can also move back and forth between already-reached stages.

Almost no models of behavior focus on repeat behaviors and habits, despite the importance of sustained change in real-life contexts. Different drivers become more or less important for repetition and habit. The intention to repeat a behavior becomes more likely when two factors converge: a positive evaluation of the previous experience ('experienced utility'), and a revised self-efficacy in comparison to what was expected[16,29]. However, merely repeating a behavior does not create habits. The creation of a habit requires the development of automaticity (the behavior is performed with low awareness, control, attention, or intention) and an association of a behavioral response with contextual cues and an experience of reward[30–32]. Some behavioral drivers that can be targeted for single actions, such as intentions, goals, or beliefs, are much less important to the formation of a habit[16,30,32]. For example, delivering a child in a health facility is a one-time behavior, for which intention, goals, and beliefs matter greatly, but exclusive breastfeeding after the child is born is close to a habit, which after initiation does not require a woman to form the intention from scratch every single time. Instead of targeting beliefs and intentions, then, restructuring of environmental cues and conscious inhibition of unwanted habits may have greater success in creating lasting habits[33,34].

Contextual drivers of behavior

Behavior emerges out of a complex system of interactions between individuals and the systems they act in. The Social-Ecological Model introduces the concept of ecosystems, which examines the dynamic ways that different layers of the social sphere influence each other[8,35,36]. Adapting the Social-Ecological Model[8,35,36], social norms and customs influence individuals in an ecosystem in several layers, and the central tenet of a norm is that deviating from it invites judgment or other negative consequences. These norms may be explicit, but they can also emerge from an implicit layer of what Practice Theory defines as

Table 1. Behavioral models surveyed, and their main advantages and limitations.

Models of behavior surveyed			
	Origin sector	**Main advantage**	**Main limitation**
Main focus: perceptual drivers			
Health Belief Model	Psychology, public health	Very widely used, wealth of data demonstrating that components explain some variance of behavior.	Neglects factors other than beliefs (biases, emotions, habits) and context/environment.
MINDSPACE Checklist	Public policy (interdisciplinary influences)	Concrete, practical checklist of evidence-based techniques to effect change across many sectors.	Focuses almost entirely on unconscious processes and corresponding nudges.
Integrative Model of Behavioral Prediction/Reasoned Action Approach/Theory of Reasoned Action	Psychology	Differentiates between different kinds of beliefs.	Context/environment is only accounted for superficially. Does not elaborate on how beliefs are formed; neglects intention–action gap (focus on intention, but intentions do not equal actions) and unconscious processes (e.g. biases).
Transtheoretical Model (Stages of Change)	Psychology	Change-as-process over time is unique component.	Evidence for six clearly delineated stages of change is weak.
Health Action Approach	Psychology	Stages of change extended to repeat behaviors.	No recognition of biases or contextual factors.
Self-determination Theory	Psychology	Differentiates between extrinsic and intrinsic motivations and names drivers for intrinsic motivation.	Focused on only one aspect of decision-making: ignores all non-motivational individual and systemic factors.
OCEAN model of Personality	Psychology	Trait-based models of personality reliably explain part of the variance in (health) behaviors.	Factors only account for part of an individual's personality, which in turn only accounts for parts of their behavior. Personality has limited predictive power for a specific behavior, but rather for patterns of behavior.
Theoretical Domains Framework	Psychology	Validated and extensive list of barriers and facilitators.	Biases and personality mostly absent.
COM-B ('capability', 'opportunity', 'motivation' and 'behavior')	Psychology	Emerging from the Theoretical Domains Framework, the first model to link different intervention and policy categories to behavioral drivers in a systematic and parsimonious way.	Limited dimensions of drivers of behavior makes the model easy to understand, but it does not provide much detail.
Fogg Behavior Model	Psychology	Similar to COM-B: behavior is understood as a mixture of motivation, ability, and prompts. Uniquely, strong focus on characteristics of contextual cues that are most effective in shifting behaviors.	Model's view of motivation and ability is simplistic.
Expected Utility Theory and Prospect Theory	Behavioral economics	Gives insight into appraisal process of a decision.	Accounts for a small subset of drivers of behavior.
Collection of cognitive biases and heuristics	Behavioral economics, psychology, neuroscience	Insight into 'automatic' and unconscious drivers of behavior.	Accounts for only one aspect of decision-making.
Evo-Eco Approach	Evolutionary biology, neuroscience	Evolutionary aspects of behavior and embodiment given due importance (e.g. disgust as a primal emotional reaction).	Views behavior as largely caused by automatic/habitual processes.
Main focus: contextual drivers			
Social-Ecological Model	Psychology	Shows the dynamic ways that different strata of the social sphere influence each other.	Does not account for perceptual drivers of behavior.
Social Cognitive Theory	Psychology	Shows how social influence can mediate some perceptual drivers.	Focuses most on self-efficacy, little emphasis on context.
Practice Theory	Sociology, anthropology	Focuses on environmental constraints on behavior.	Neglects individuals, focus on theoretical level rather than testing components' explanatory value.
Diffusion of Innovations Theory	Communication studies/ sociology	Clear guidance on techniques to reach different segments of a population to adopt a novel behavior.	Segments individuals in a specific way (how receptive they are to an innovation), does not account for other environmental and cognitive factors driving decision-making.

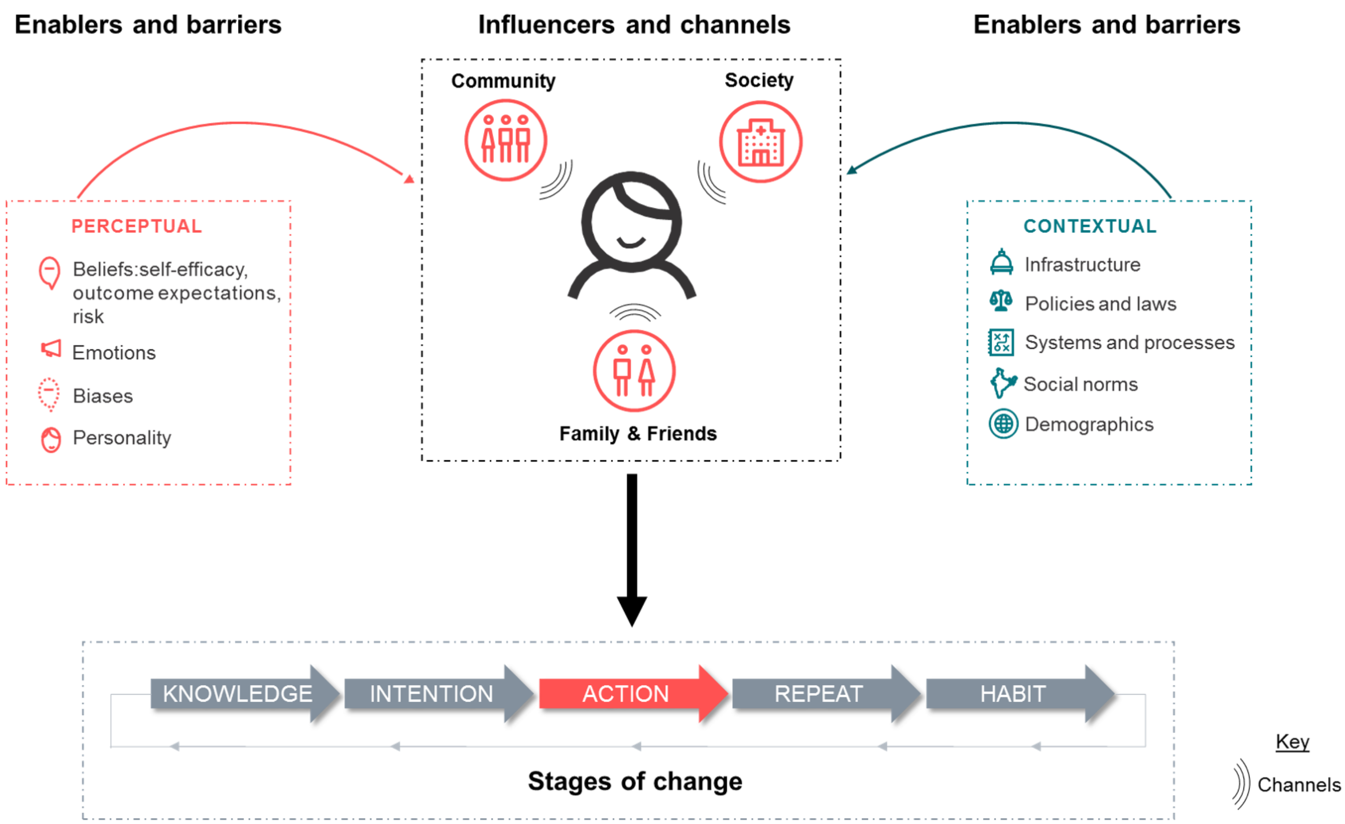

Figure 1. The CUBES behavioral framework. Contextual and perceptual drivers combine to act as enablers and barriers along an individual's path from knowledge – encompassing awareness and skills – to intention (or motivation to act) and action and beyond. Layers of influencers can affect these drivers and reach an individual through various channels.

'shared cultural knowledge' that expresses itself as routines and habits[7,37].

Structural factors are further contextual drivers that may shape and constrain perception and action. These aspects are usually not fleshed out in behavioral models. However, unless these constraints are removed, interventions acting on perceptual drivers will not allow for the target behavior to occur. We propose differentiating between infrastructure, policies and laws, and systems and processes, all of which vary strongly with their respective context. For example, infrastructure drivers may include availability and condition of roads leading to a health facility, equipment to perform a test, or seeds for farmers to use. Policies and laws constraining behavior can exist on several levels, from national laws to facility-level guidelines. Systems and processes could mean supervision, training, feedback, or incentive systems under which healthcare providers operate in a facility, or the tools teachers have available to plan lessons and receive feedback on their tuition. Finally, demographic factors such as age or education level act as contextual constraints on behavior.

Perceptual drivers of behavior

Multiple models recognize that behavior is also shaped by perceptual drivers that, together with contextual drivers, combine to make a target behavior more or less likely.

Biases and heuristics. Most aspects of an individual's cognition and behavior are influenced by 'automatic', often unconscious, mental shortcuts or rules of thumb (heuristics) that can bias decision-making. Many biases can be difficult to change, but knowing about them makes it possible to construct environments where the best decisions are the easiest. A typical example is setting the default option of a pension program to 'opt-out' instead of 'opt-in': staying in a pension program after auto-enrollment is easier than making the effort to join in the first place[38,39]. Examples of heuristics and biases are the optimism, confirmation, and availability biases, anchoring-and-adjustment, hyperbolic discounting heuristics, and the status quo and representative biases[17,18,38–44]. MINDSPACE, an influential checklist for behavior change developed in part by UK government agencies[20,45], is an example of using biases (such as a bias to choose the default option) as tools to design interventions.

Beliefs. Beliefs are formed by learned experience, differentiating them from biases (which humans share to varying degrees as a result of how our brains evolved). In the Integrative Model of Behavioral Prediction/Reasoned Action Approach, beliefs about what outcome can be expected from a behavior (called 'attitudes' in that model), normative beliefs (how others will judge a behavior), and beliefs about the extent of one's control over the behavior (self-efficacy) all influence intention, which is seen as the main driver of behavior[19,46,47]. Experiments have shown that some beliefs predict behavior better than others. For example, perceived control (self-efficacy) and beliefs about a behavior's outcome are better predictors of behavior than normative beliefs[48,49]. Indeed, self-efficacy beliefs have emerged as a strong predictor of behavior across other models, such as Social Cognitive Theory, the Health Belief Model, and the Transtheoretical Model. The strong influence of self-efficacy on behavior has been shown experimentally in many studies[50–54].

Outcome expectations are another example of a belief. An individual appraises a potential behavior by weighing perceived costs and perceived benefits, which may be emotional ('How will the outcome of the behavior make me feel?'), social ('How will others judge me?'), or functional ('How does this help or hurt me?'). Outcome expectations, together with the perceived severity of an outcome, the perceived susceptibility to that risk, and self-efficacy, are central to one of the most widely-used models of behavior, the Health Belief Model[14,24,55]. Evidence from several systematic reviews shows that increasing perceived benefits and decreasing perceived costs to a behavior will be most likely to cause an individual to engage in the target behavior[10,24,56,57].

Emotion. The experience of emotion (affect) also drives behavior. Affect arguably colors all perception and powerfully shapes decision-making[20]. While affect is used as a distinct driver of behavior in the MINDSPACE checklist[20], the COM-B framework[22] includes it as a rapid, automatic component of motivation. One example of targeting affect to drive motivation to engage in a new behavior can be seen in an intervention promoting soap use in Ghana: education around the benefits of soap did little to drive up its popularity, but emphasizing the feeling of disgust from 'dirty hands' resulted in significantly increased soap use[58]. This emphasis on automatic emotional responses, such as disgust or a desire to conform with others, is a key component of the Evo-Eco behavioral model[59].

Personality. Personality traits are not often included in models of behavior—not even in the Theoretical Domains Framework, which includes the perhaps most comprehensive list of barriers and facilitators[60] —but they can strongly influence an individual's propensity to engage in and maintain health behaviors[61,62]. Current applications of personality models to behavior prediction focus on aggregates of behavior, or behavioral patterns of behavior such as going to check-ups and regular physical activity (preventive health behaviors). Such models often do not attempt to predict single instances of behavior, which has been shown to be much less reliable[61]. Currently, the dominant personality model with a large evidence base behind it is the so-called Five-Factor Model[12,62–64] with the five broad traits of openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism ('OCEAN'). Conscientiousness appears to be an especially strong predictor of behavior patterns, such as sticking to preventive health routines[61,62].

Influencers and channels

Following from the Social-Ecological Model[8,35,36], influencers surround an individual in layers. Family and friends or peer groups are the closest layer to the individual. The next layer usually consists of relationships in the community, for example in workplaces, schools, and neighborhoods. The most distant layer is the larger social context. Influencers can reach individuals either directly or at scale via various media channels, which is important information to determine how to deliver interventions. For example, female self-help groups can serve as a channel for rural women in India to reinforce or change social norms relating to a certain target behavior. For an intervention to work, the content, the type of influencer, and the channel through which they reach individuals must be identified as relevant to the target individuals.

Interaction between the 'building blocks' of CUBES

Depending on the behavior, all the drivers mentioned above will be relevant to varying degrees to any one individual, and their combination will result in a larger or smaller tendency to act. The elements of CUBES influence each other deeply. For instance, punitive supervision (systems and processes) by medical-officers-in-charge (influencers) in a health facility might lead nurses on the receiving end to experience high anxiety (emotions), and to beliefs that trying hard will not result in any benefit to them (outcome expectation belief). It is important for programs to disentangle these components, even when they influence each other, because this determines what kind of intervention is best placed at what level.

Therefore, the enablers and barriers in the different 'building blocks' of CUBES can be seen as a checklist for programs that they can utilize to design effective interventions.

Models of behavior that simplify to the point of 'motivation' or 'ability' (such as COM-B or the Fogg model) are simpler, but also less actionable for programs on the ground. For example, if a study finds a lack of motivation to go to the doctor despite having symptoms indicative of tuberculosis, this alone is not actionable by an intervention. Instead, programs need to know where potential patients are on the knowledge–intention spectrum (they have clearly not yet taken action), whether they perceive their symptoms and the disease as a danger to their and others' health (risk perceptions), whether they think going for a check-up will actually alleviate symptoms (outcome expectations), whether they feel able to skip work and other responsibilities to attend an appointment (self-efficacy), whether appropriate facilities are even available, affordable, and accessible (structural factors), and whether there is stigma involved in seeking care (social norms). All these components would feed into the concepts of 'motivation' or 'ability', but require very different interventions to effect change.

Once CUBES has been used to understand and categorize existing evidence, evidence gaps can be closed in a focused way with primary research. Not all components of CUBES are best captured and intervened on in the same way. In the following sections, we introduce a method mix designed to identify a comprehensive set of drivers.

Methods of measuring drivers

Qualitative interviews, focus groups, and quantitative surveys are some of the most common methods of insight generation in the global development. These strategies complement each other: qualitative methods are best suited to exploration and capturing nuances, whereas quantitative methods are indispensable for discovering patterns and weighing the relative importance of different drivers, which is essential for developing interventions that address the barriers that matter. Here, we propose two overarching considerations for programs to add value to the methods used.

First, existing methods can be improved by using the CUBES framework as a checklist against survey or discussion-guide items, to check whether a comprehensive set of enablers, barriers, influencers, and stages is captured. Too often, methods such as quantitative surveys remain at the level of measuring practices (or '*what*' data) and demographics, at the expense of the '*why*', or perceptual and contextual drivers of the target behaviors.

Second, programs could benefit from expanding their own toolkit of methods by selecting approaches to investigating behavior from a variety of sectors. The right method will depend on what type of data needs to be captured, and whether the purpose of research is exploration or testing specific hypotheses. A method mix can also help counteract any weaknesses of individual methods.

Choosing the right type of method for different stages of research

Research methods can be divided into descriptive, experimental, and simulated approaches; the last two are relatively underused in global development. Whether qualitative or quantitative, *descriptive* methods such as interviews or observation aim to describe and explain behavior without testing the effects of manipulating variables systematically. While they can explore 'how' and 'why' questions (and are therefore often called 'exploratory'), they do not have the ability to systematically relate the effect of change to outcomes ('confirmatory'). *Experimental* approaches can be used to test hypotheses and find causal relationships by systematically varying variables and testing their effects on outcomes. However, they lack the ability to survey a broad spectrum of factors at once that exploratory methods provide. Finally, *simulated* methods can assess cause–effect relationships in a virtual environment when experimental field methods are not possible or are too complex, but they rely on many assumptions to construct the simulation.

Expanding the descriptive toolkit

Journey mapping. Developed and primarily used in market research[65,66], journey mapping systematically tracks people's experiences and interactions with a product, service, or life event over time, as people form beliefs about the product or event and make decisions, perhaps via influencers, to interact with or avoid it. This method is especially well-suited to get a sense of stages of change. Journey mapping can also help form hypotheses about segments of customers who share distinct characteristics in order to target them with bespoke messages via different channels. In addition, it can be useful in generating hypotheses about underlying behavioral drivers that can then be tested further.

Journey mapping uses many different techniques to collect data, including one-on-one qualitative interviews, focus groups, ethnography, web analytics, customer reports via apps, and (qualitative) network mapping[65,67]. Below, we show in a case study how journey mapping was successfully integrated in a program understanding decisions around voluntary medical male circumcision[5].

Observation. Observation is a versatile and routinely used tool in global development to collect data about what people do, the context that surrounds them, and how they interact with processes, objects, or each other[68]. The spectrum of observation techniques ranges from researchers interacting closely with communities (*participant observation*), to covert or overt '*natural observation*' without participation, to *controlled observation*, where procedures are highly standardized. Measuring time and frequency of practices of medical residents in hospitals[69], or of nurse practices in hospitals in India[25], are typical examples of controlled observation. Such *time-and-motion* studies observe the time taken and actions of participants executing distinct components of a process. Observation is also a tool to measure contextual drivers such as infrastructure or processes. For example, *facility and infrastructure audits* commonly combine observations with interviews of key stakeholders to track characteristics such as hospital staff coverage, equipment availability, or communication tools[70].

To get a holistic sense of contextual drivers, observation can track more than behaviors or supplies. Instead, immersive observations in the participants' natural environment could be structured to assess the set of contextual enablers and barriers outlined in the CUBES framework. We call this approach '*structured immersive observation*'. For example, in a recent observational study on nurses in healthcare facilities, we measured key contextual dimensions as follows. First, we assessed facility infrastructure available to nurses: whether equipment and drugs required for routine tests were available at the time of testing, staff coverage throughout the period of observation, availability of beds, water, and electricity, and transport options for patients to be referred. Second, we assessed systems and practices, such as interactions with and feedback from other staff, time spent with patients and tests performed, documentation systems and job aids, and training records. Third, we assessed community norms by observing community interactions and communications with the nurse.

Enhanced quantitative surveys. Quantitative surveys are a critical tool to obtain insights on many enablers and barriers to behavior simultaneously and at scale, and consequently are a mainstay of global development research. Survey design is a broad field with a large array of approaches. Here, we recommend three key techniques that can enhance the design of quantitative surveys to measure potential behavioral drivers. Survey questions can be structured to account for as many components of

CUBES as possible. In sensitive contexts, surveys can also be enhanced to counteract respondent biases. Finally, while not within the scope of this article, programs would benefit from leveraging quantitative data for insight beyond descriptive analyses: for example, population segments can be found for targeted intervention design[5].

Driver-structured surveys. Programs can use the CUBES framework as a checklist to assess whether a survey captures the range of potential contextual and perceptual enablers and barriers, influencers and channels, and stages of change relating to behaviors of interest. We often find surveys only measure a narrow set of drivers, which presents the opportunity to generate a more holistic view of behaviors of interest and the system. This can also be enhanced by adapting existing validated tools, such as scales testing personality or self-efficacy (see below).

Standardized scales. A simple and high-yield modification to quantitative surveys is the inclusion of previously validated and standardized rating scales relating to CUBES perceptual drivers. Standardized scales that test specific cognitive processes include the Risk Propensity Scale[71] and the ten-item General Self-Efficacy Scale[72]. Types of emotions and their felt strength have also been widely measured with graphical rating scales, such as the Self-Assessment Manikin[73]. All these scales can be flexibly adapted to specific contexts. For example, in a study investigating women's propensity to engage in breast cancer prevention, self-efficacy was asked in two items: 'the extent to which participants were confident that they could conduct breast self-exams every month; and when they conducted a breast self-exam, how confident they were in their ability to identify a "lump that needs medical attention."'[74].

Personality tests, widely used in the private sector, also use standardized rating scales. Many studies show some predictive value of the OCEAN model's 'Big Five' personality traits on health behaviors[61], especially conscientiousness[62]. Questionnaire designers can tap a large number of validated instruments to test OCEAN components, such as the public-domain International Personality Item Pool[75].

All standardized scales have the advantage of using previously validated instruments, and that individual differences can be captured with relatively little effort. A limitation is that, like all self-reports, such tests are susceptible to reporting bias, since participants can deduce or guess socially desired responses (those that the respondent thinks will make them appear in a favorable light). Responses may therefore be compatible with the participants' sense of self rather than their actual behavior.

Informal confidential voting interviews and polling-booth surveys. Methods that stress anonymity and confidentiality, such as polling-booth surveys (PBS) and the Informal Confidential Voting Interview (ICVI) approach, can be used to probe sensitive topics, as they counteract social desirability bias. The ICVI consists of a one-on-one interview followed by self-completion methods[76]. Similarly, PBS collects feedback from a group of people who respond anonymously and independently through a ballot box. Comparison of one-on-one interviews with PBS[77] and with ICVI[78] on sexual risk behavior in Indian men and women demonstrated their value, as more risky behaviors were reported with each method.

Standardized patients. Standardized patients (SPs) are people trained to play the role of a patient with certain medical and personal characteristics, who interact with healthcare providers in a realistic setting. SPs are comparable to 'mystery shoppers' in consumer research[79], in that the healthcare professional does not know the patient is not real[80]. In other scenarios, both parties know about the setup[81]. The SP method has been applied to investigate healthcare provider behavior in many contexts, such as prescription practices of pharmacists[79]. It has also been used to assess how doctors communicate with their patients, such as how surgeons disclose medical errors to patients[81]. The SP approach can compare expected with actual behaviors, and analyze communication, such as the content of advice given to patients (80). While the method on its own is very well suited to capture practices and contextual drivers such as infrastructure and processes (of patient interaction), other drivers, such as beliefs or biases, are less accessible to investigation.

Social network analysis. Social network analysis (SNA) maps relationships between people, organisms, groups, or devices[82]. When analyzing behavior, SNA can be an excellent tool to describe which influencers and which channels are most important to transmit certain norms and information. SNA can be both qualitative or quantitative. Data can be generated from surveys, ethnography, or observation, or mined from existing resources such as GPS coordinates or twitter messages[83,84]. Many field studies have used SNA to focus on where in the system to intervene. For example, in Uganda, researchers mapped the process of obtaining a diagnosis for tuberculosis through provider and patient networks, and the steps where delays were most common could be identified[85]. Ultimately, SNA is a flexible and versatile method, but specifically focuses on identifying centers of influence.

Leveraging 'passive' datasets. As in the SNA example, insight can also be generated from leveraging 'passive' datasets, generated for a different original purpose, without direct interaction with or observation of respondents. Examples are information obtained from GPS, satellites, and sensor systems, as well as other databases. To investigate contextual drivers, satellite images can map physical conditions of the built environment, which can then be related to behaviors and drivers from other datasets[86]. For some audiences, social media data can be an appropriate source of aggregate estimates of positive or negative sentiment[87]. More analog data sources can also help generate insight, as in an analysis of Kenyan newspaper articles about voluntary medical male circumcision, which provided insight on the types of risks that were presented to readers[88].

Assessing decision drivers through '*in vitro*' experiments

We propose that programs can benefit from '*in vitro*' experimental methods before testing specific interventions in lengthy trials in the field. Experimental methods that track the decisions participants make in laboratory-like conditions serve several purposes: programs can systematically change and test enablers and barriers to behavior, predict behaviors in response to specific interventions, determine those features of a service or product that are most likely to align with the customers, and forecast the market size of a product or features based on predicted behaviors. All these factors narrow down the potential characteristics of an intervention to be tested in the field, and ultimately make the design of effective interventions more probable.

Discrete choice experiments. Discrete choice experiments (DCE), extensively used in market research, uncover preferences and value attribution from the choices that participants make, rather than from the participant disclosing them. They are a powerful tool to predict behaviors '*in vitro*', assess which features of a product or message are most important to the customer, and to forecast market shares of products. DCE have been shown to be predictive of health behaviors[89]. For feature selection, participants are typically shown multiple iterations of sets of products with varying features. In each trial, the participant picks one option. For example, a discrete choice experiment in South Africa evaluated which characteristics of HIV prevention products, such as the method of use, or the protection against diseases other than HIV, would be most valued by participants[90]. From participant choices, a model can be built showing which level and which combinations of a product's features predominantly drive decisions, where the tipping points of certain preferences lie, and forecasting product market share. DCEs can also be used to test various 'what-if' scenarios, and results can then be used as a funnel to select the most promising attributes for a field intervention. Simulated test marketing is a related concept, in which the consumer is asked to make choices in a realistic environment, with similar systematic manipulation of test variables.

Decision games. Purely quantitative DCE approaches are mostly used to cycle through permutations of features and types of products or interventions. However, a related 'decision game' method mixing quantitative with qualitative elements can help investigate which behavioral drivers most influence choices made by participants. In a recent study with healthcare providers in Uttar Pradesh, India, we used such a group 'decision game': participants were given a set of scenarios, each with a set of response options, and were asked to choose the option they thought other participants would select (unpublished reports). Response options coded for different behavioral drivers, and participants were later qualitatively probed on their choices. For example, nurses were asked which nurse in a scenario was likely to be most stressed: the one working in an understaffed facility (coding for infrastructure-staffing), the one dealing with demands from patient family members (influencers-community), or the one facing constant scrutiny and accountability from supervisors (systems and processes – supervision).

Implicit attitude tests. Few research methods are suitable to measure enablers and barriers that respondents cannot or do not want to report. Over the last two decades, experimental psychology has developed a battery of experimental approaches to measure 'implicit' biases, or biases that are inaccessible to conscious awareness and self-report, but nevertheless influence behavior. The underlying concept of implicit attitude tests is that our brains perform unconscious evaluations of concepts, people, and objects, which have arisen from past experiences and cannot be measured by explicit questioning.

The most widely-used implicit attitude test is the Implicit Association Test (IAT). This test is based on the concept that participants can perform a task more quickly when they see two concepts as related than when they do not associate them with each other[91]. IATs have been used to measure a plethora of social stereotypes, such as gender and racial biases[ii], for instance in the Democratic Republic of Congo[92]. In market research, the IAT has been used to gauge consumer attitudes toward different products[93]. The predictive value of implicit attitude tests on behavior is still under debate[94,95]. For this reason, and because IATs can only test a small set of associations within a test that requires training participants, we only recommend implicit tests when a behavior is likely to be influenced by a specific, deep-seated bias that respondents are unlikely to report.

Using simulations to model 'what-if' scenarios

Simulations have the unique advantage that complex 'what-if' scenarios can be explored at the push of a button, which can be used to supplement and inform data collection or to optimize interventions. Through the construction of 'virtual worlds', mathematical models can simulate the impact of implementing certain interventions, or of targeting interventions to specific sub-groups. They can also generate hypotheses on what a likely driver for behavior might be. As an example, agent-based models have been used to simulate the large-scale effects that emerge from the actions of many single agents, such as the spread of disease or of social norms and beliefs[96,97]. Similarly, Bayesian cognitive mapping builds probabilistic models of the likelihood that agents make certain decisions[98]. However, simulations are only as good as the model and assumptions that underlie them. The relevant and correct starting parameters must be chosen with caution, which includes a degree of subjectivity[99], and generalizations from a model based on specific assumptions are limited. Table 2 summarizes the approaches discussed, and their strengths and weaknesses.

Choosing the right method at the right time for the right purpose

The goal of any program is to implement successful interventions in the field. Figure 2 depicts the process of setting the research agenda, generating insights, and designing and optimizing interventions that programs can use, depending on the knowledge level at the start of the research process. First, when programs define a target behavior to be changed, they can evaluate existing evidence against the components of the CUBES framework. This can be done either from existing literature or analyzed from datasets, or both. To directly choose an intervention that works, programs must already have narrowed down specific drivers to intervene on. This may be the case in a data-rich environment; in other cases, political or resource constraints limit what is

[ii] Project Implicit Research Group. Project Implicit 2011. Available from: https://implicit.harvard.edu/implicit/ethics.html.

Table 2. Overview of enhanced and novel insight generation methods as part of the CUBES toolkit.

Method	Primary insight gain	Most testable CUBES components	Method type	Advantages	Disadvantages
Descriptive					
Journey mapping	Tracking experiences and influencers over time	Stages of change, beliefs, emotions, influencers and channels	Qualitative	Mapping experience, influencers drivers over time	Self-report
Observation ▪ Time-and-motion ▪ Infrastructure audits ▪ 'Structured immersive observation' (SIO)	Systematically tracking practices/ duration (time and motion), infrastructure and supplies (audits), and CUBES-structured contextual drivers (SIO)	Contextual drivers: structural, systems and processes; behaviors observed	Quantitative	Behaviors and contextual drivers and barriers can be measured in their natural environment, in a standardized and replicable way	Observed participants and researchers are prone to behavioral or recording biases, respectively.
Enhanced surveys					
▪ Driver-structured surveys	Using CUBES as checklist aids systematic capture of enablers, barriers, influencers, and stages of change	All	Quantitative	Holistic overview of all potential drivers possible in one dataset per respondent	Not all drivers are equally well captured by self-report
▪ Informal confidential voting interview (ICVI), polling-booth surveys	Adding anonymized components encourages responses on sensitive issues	Social norms, beliefs	Quantitative, ICVI also qualitative	Greater disclosure on sensitive issues	Yes/no response format leaves no room to explore; anonymous data can only be analyzed in aggregate
▪ Standardized scales	Testing perceptual drivers with validated, standardized tools (e.g. self-efficacy, risk propensity, personality)	Beliefs, personality	Quantitative	Ready-made aids to assessing perceptual drivers and barriers	Prone to self-report bias
Standardized patients	Tracking behaviors, context, and interactions through simulated 'patients' with a set of standardized characteristics	Contextual drivers: structural, systems and processes; behaviors observed	Quantitative, qualitative components	Standardization allows for comparability, realistic setting and covert data collection for realism	On its own, is mostly limited to 'what' data and cannot explore drivers for practices.
Social network analysis	Revealing direction and strength of relationships in a system	Influencers, social norms	Qualitative or quantitative	Versatile (qualitative or quantitative), useable for networks of any size and type, unique method of identifying influential targets for potential intervention	Network modeling can only investigate limited drivers and barriers in one network.
Leveraging 'passive' datasets	Generating insights from sensor, mobile phone, satellite, GPS, social media, and other databases, with no direct customer interaction	Different, depending on dataset	Quantitative	Large-scale existing datasets can be tapped and integrated with other research methods, 'bird's-eye' view of context possible	Passive nature means no opportunity to probe; existing datasets may not focus on key customer groups
'*In vitro*' experimental					
Discrete choice experiments	Participants make repeated choices between a set of options whose attributes are systematically varied, in order to uncover which attributes are most important	All, least useful for biases	Quantitative	Quick to develop, test, and analyze. Participants do not have to explain 'reasons why', which are inferred from choices	Correlation of hypothetical with real-world choices is difficult to predict. Providing response options that clearly represent distinct drivers and barriers is not trivial
Decision games	Gamified, social experiment version of a discrete choice experiment	All, least useful for biases	Quantitative and/or qualitative	Gamification increases engagement, asking about what other participants select instead of own choices circumvents some respondent biases	Same as discrete choice experiments; qualitative approach is difficult to interpret
Implicit attitude tests	Using reaction time in response to tasks and other measurements to determine whether participant sees concepts as related or not	Biases	Quantitative	Unique method to assess strong biases inaccessible to self-report or observation	Method not well tested in low-resource settings, correlation of output and behavior not obvious, each test can only test a limited number of associations
Simulated					
'What-if' simulations	Modeling simulated decision-making or outcomes in response to changing parameters in complex systems	All can be simulated	n/a	Unlimited permutations of changes ('what-if scenarios') in a complex system can be modelled	Any model will only be as good as the input data (which does require field-level input), highly specialized skills required

testable in the field. In such cases, primary research may not be required or appropriate. On the other end of the spectrum, a program might know what it wants to change—for example, to increase the uptake of modern methods of contraception—but have little systematic knowledge of the types of drivers that may be involved. In this case, exploratory research is warranted ('insight generation' in Figure 2). We have found that a quantitative survey often provides the best practical balance between assessing many components of CUBES at scale. Either before a survey (to inform its design) and/or after (to dive deeper into specific findings), specific descriptive or experimental methods offer particular strengths assessing specific CUBES components and can supplement a survey (Table 2).

Descriptive methods can be qualitative or quantitative, and many can take both forms. To choose a qualitative or quantitative focus, programs can consider whether the freedom to explore limited aspects in depth is most important (which means a greater focus on qualitative research), or representativeness at scale and the relative likely impact of each driver (which points to quantitative methods). Often, programs use preliminary qualitative research to the inform the design of quantitative

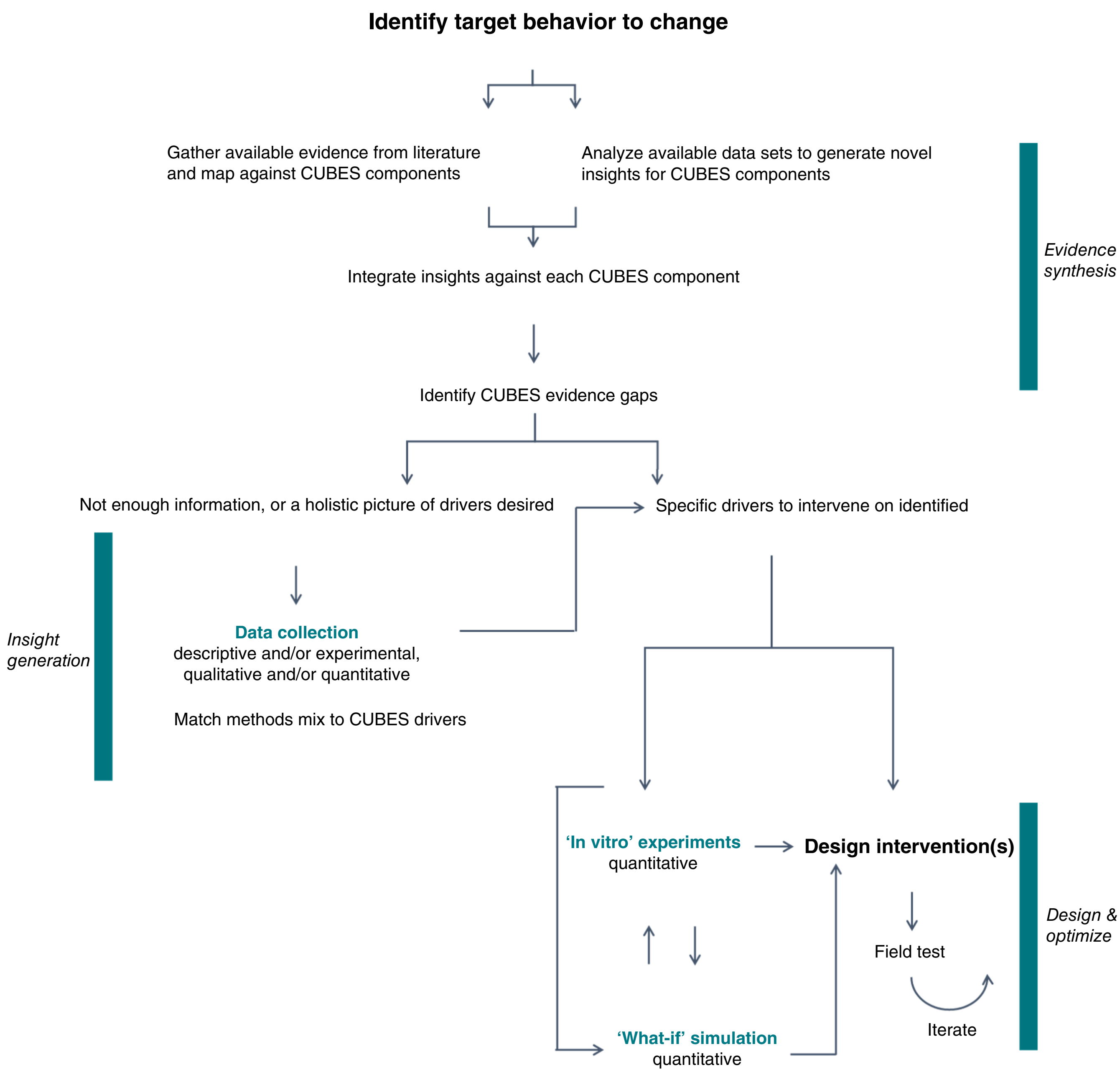

Figure 2. Decision aid for choosing the right research approach at the right time, for the right purpose.

research, especially in field settings where quantitative research is conducted in person and therefore is expensive and time-consuming, whereas small-sample qualitative research is less resource-intensive. The freer structure of qualitative research also typically allows for follow-up questions and clarification. However, we recommend that this order not be followed by default, but rather examined on a case-by-case basis. In our experience, qualitative research can sometimes divert resources from quantitative research, which can be wasteful if results from a smaller sample cannot be generalized. Instead, qualitative and quantitative research can also be run in parallel, with a complementary focus on different drivers; pockets of qualitative components can be mixed into qualitative research; or qualitative back-checks can be conducted after quantitative research.

In global development, most programs tend to focus on descriptive methods for insight generation, followed by field implementation, where randomized controlled trials assess the effectiveness of an intervention. In addition to preliminary evidence synthesis and using a flexible methods toolkit for specific deep-dives, we argue that this approach misses a key step, namely '*in vitro*' experimental methods. These methods can narrow down the many potential hypotheses emerging from exploratory research, so that only the enablers and barriers likely to be most impactful are ultimately tested in the field. This optimization can be used to choose between different types of interventions (such as monetary incentives versus more information on risks and benefits), as well as the components of a specific interventions (such as the magnitude of incentives likely to be most effective). If rich descriptive data is already available, and so a limited set of specific hypotheses around limited drivers can be formed from the outset, programs can directly skip to this step. This step can also be done in parallel with exploratory research, if evidence is strong on specific drivers but weak on others and a holistic picture is desired. As detailed above, purely simulation-based methods can be of use here to model the effect of many different changes. As a result of this step, field testing will be based on much stronger evidence.

Applications of the cubes toolkit: case study

Designing interventions to increase the uptake of voluntary medical male circumcision (VMMC)

In Figure 3, we briefly outline the methodological choices made to investigate enablers and barriers to uptake of voluntary medical male circumcision (VMMC) in Zambia and Zimbabwe. VMMC is a highly cost-effective intervention for preventing HIV acquisition that is being scaled up in eastern and southern Africa[5,100]. The achievement of the program's ambitious targets necessitated shifting the behavior of many men in the community who either did not consider circumcision, or if they did, did not take action. Therefore, the program needed to understand the multiple interacting factors that facilitate or inhibit men's decision to get circumcised, and to test interventions to address those factors. A synthesis of previous studies revealed a variety of existing insights on many behavioral drivers, such as concerns around pain, complications, or cost, as well as patterns of influencers[101]. Analysis of a large quantitative survey further showed a relatively small awareness-intention gap, but a large drop from intention to action[102] A total of 64% of men intended to get circumcised, but only 11% did. However, a holistic view assessing the prevalence of each of these, and other, drivers and their relative strength was lacking, and most research was either small-scale or qualitative[5]. The existing studies could not answer why there was a strong intention-to-action gap. Also, the research did not examine or reveal the heterogeneity among men—the fact that a given enabler or barrier may be important to one man, while not as relevant to another.

A broad set of CUBES drivers therefore needed to be captured at scale to assess the relative importance of each driver in a single dataset. However, as the stages of change appeared to be of primary importance, journey mapping and qualitative decision games were first used to understand the stages of change and associated beliefs and influencers at each stage in more detail[5]. In summary, qualitative research pointed to a much more nuanced picture. Men develop positive as well as negative beliefs, influenced by individuals around them, as they move through the various stages of change. These competing beliefs and associated emotions move men towards or away from the decision of getting circumcised in distinct stages of change. For example, beliefs in the early stages include "VMMC protects myself and my partner from STIs" (positive) and "the procedure is painful". As men move to consider undergoing the procedure, there emerges a strong conflict between the positive beliefs and negative ones, such as circumcision threatening self-identity, leading to distrust between the man and his female partner, and the perceived long healing time. The conflict between emotions such as shame, distrust, and fear and perceived potential benefits move men into a state of cognitive dissonance and stall them from taking-action. However, it was also clear that not all men held all beliefs equally strongly, or were equally subject to the same type and strength of influencers. These insights informed the design of a large-scale quantitative survey investigating CUBES components comprehensively. Having this dataset available then allowed us to segment men on the enablers and barriers to VMMC, so that messaging interventions could be designed targeting the most important drivers for each segment. For instance, among the six segments found in Zimbabwe, Embarrassed Rejecters had mostly negative beliefs about VMMC, as well as fears and concerns regarding the procedure, and had little social support[5]. However, they did not lack the knowledge that strongly characterized another segment, Neophytes. Accordingly, segment-specific interventions, including messaging through front-line workers or media campaigns, could be developed. In addition, the roll-out of circumcision devices as an intervention to improve the uptake of circumcision was an important consideration for programs. However, it was unclear what the potential demand and market share of these devices would be. The demand for different devices to carry out the VMMC procedure was forecast using simulated test marketing, a technique related to discrete choice experiments, so that the right devices could be marketed with the right message to the right people[100]. Interventions designed based on this research are currently being piloted at national scale in Zambia and Zimbabwe.

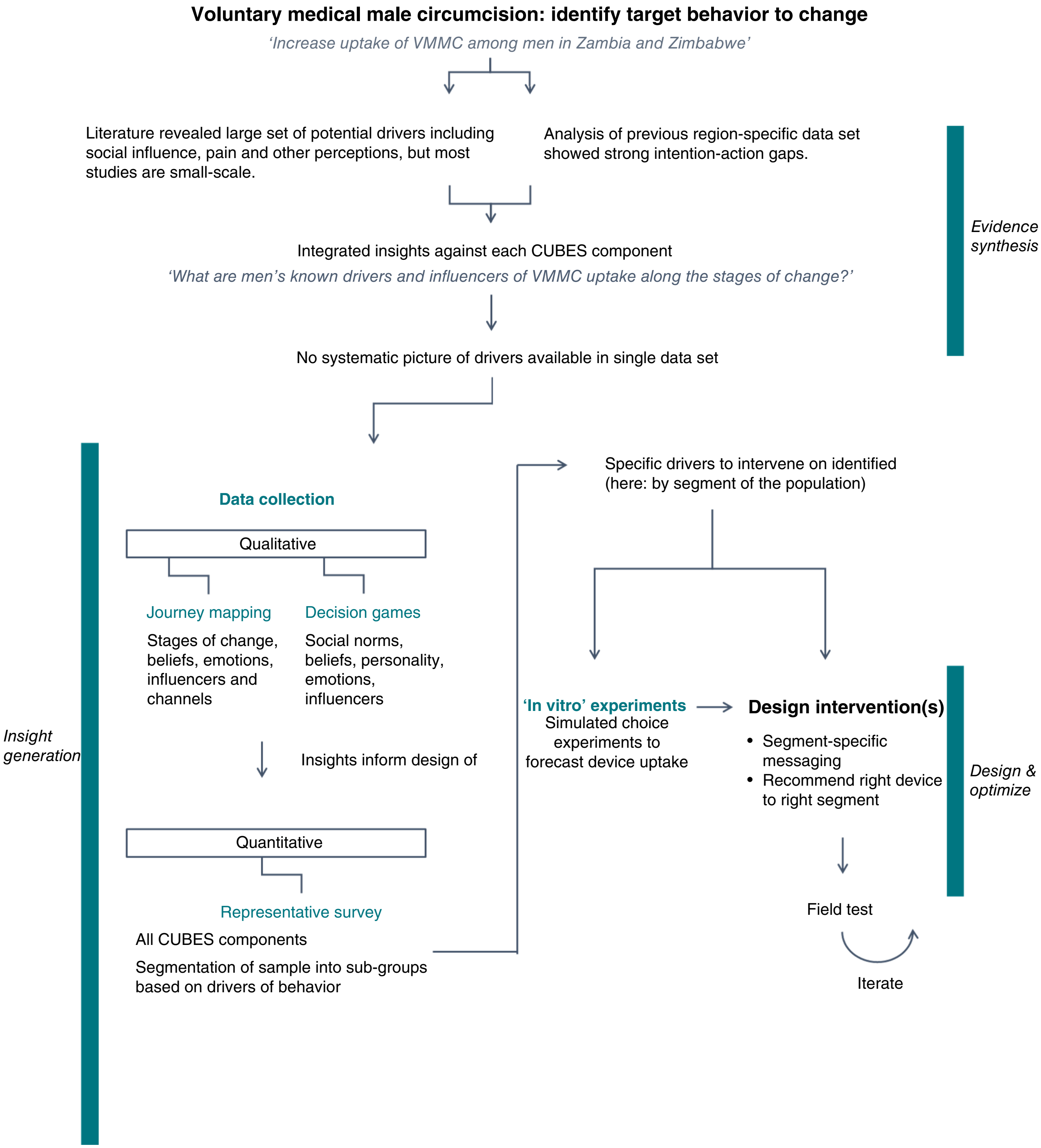

Figure 3. Process of evidence evaluation, insight generation, and intervention design and optimization in a VMMC program.

Discussion

Effective interventions to drive key outcomes are sorely needed in global development and many other sectors. In this paper, we aim to help programs arrive at an effective portfolio of interventions in two ways.

First, to effectively design interventions that change target behaviors, we introduce a novel and practical framework of behavior, CUBES, to help programs categorize and understand the barriers and enablers, influencers and stages of behavior change (Figure 1 and Table 2). CUBES synthesizes widely validated evidence across psychology, behavioral economics, market research, and sociology[7,8,11–14,16,19,20,22,34,38,45,48,53,72], and presents its building blocks with a view to actionability by programs. CUBES provides a checklist for programs to systematically assess what is already known about drivers of a target behavior, where novel research is most needed in order to design actionable levers of change, and, after closing evidence gaps, where interventions could focus.

Second, not every type of driver is best measured in the same way. We therefore curate a set of descriptive, experimental, and simulation approaches across sectors, and advocate for a method mix tailored to the gaps in knowledge in a given program (Figure 2 and Table 2). Some approaches, such as different types of observation and self-report, are already well-established in global development, but using CUBES to structure the components of insight generation ensures that programs can design tools in a systematic way, ultimately saving time and money. For example, quantitative surveys would benefit from the selective incorporation of validated scales to measure specific drivers, or from ways to encourage participants to respond to sensitive topics with greater fidelity. Other methods are well-used in other sectors such as market research, experimental psychology, and decision sciences, and programs could benefit from them for specific purposes. For instance, discrete choice experiments and decision games provide an experimental way to systematically vary and identify key enablers and barriers before testing interventions in the field. Implicit attitude test can be considered as a complementary method to self-report when strong biases are presumed to be at play, and simulations modeling complex systems provide programs with a way to estimate the importance and interaction of multiple drivers, as well as test 'what-if' scenarios. We demonstrate the process of choosing methods for specific purposes in a case study on voluntary medical male circumcision uptake (Figure 3).

Once data has been collected, the CUBES framework can again be used to structure findings and highlight the specific content and potential targets of an intervention. In a complex real-world setting, not all interventions to change behavior will be successful. However, knowing which drivers of behavior are important in a given situation can narrow the options for choosing suitable intervention types. For example, if most barriers to a given behavior are found to be unconscious biases or structural factors, then education around the benefits of an action is unlikely to lead to a different outcome; whereas if misinformed beliefs around outcomes are the main barrier, education could be an effective lever. We previously showed that psycho-behavioral segmentation can be a powerful method for finely targeting interventions beyond a one-size-fits-all approach[5,6]. In the voluntary medical male circumcision program, we used the quantitative survey data to segment men on what drove them toward or away from the procedure, and could therefore tailor interventions specifically to each segment[5]. Previously, Michie *et al.*[103] provided a comprehensive overview of 93 behavior change techniques, from social comparison to incentives, feedback on behavior, prompts, and goal setting[103], as well as a more high-level categorization of nine types intervention from education to training, incentivization, and environmental restructuring[22]. While identifying mechanisms of actions for interventions remains a work in progress[104], these categories can serve as decision aids and an overview of options to programs. The Fogg Model of behavior, as well as the MINDSPACE checklist, also provide a useful classification of what characterizes effective prompts that can increase motivation and ability[20,21]. However, sound and actionable segmentation can only be performed on large-scale quantitative datasets, which is a consideration for the method mix chosen.

After designing interventions to match key barriers found, whether at the population or segment level, we also recommend optimizing and 'funneling' interventions from a large pool of potential options down to a narrow set that can be thoroughly evaluated in the field while maximizing the likelihood of success (Figure 2). Discrete choice experiments can help programs choose between types of interventions and their components at the design stage. Even at the field test stage, factorial designs could test sets of interventions more efficiently, such as different magnitudes of monetary incentives, a distinct messaging component in each, and a different channel through which they are deployed. Recently, this approach has been refined in the Multiphase Optimization Strategy (MOST) for determining the best set of intervention components[105].

CUBES and the methods toolkit proposed here have several limitations. Frameworks of behavior in general have an 'evaluation problem', as it is not feasible to directly compare the insights generated *de novo* using many different behavioral frameworks with very different components. The true test of time will lie in whether programs judge CUBES to be useful in increasing intervention fit, as well as effectiveness at reaching target outcomes, as we have found in our own programs. A second limitation is that it is unlikely that any one program will be able to draw on expertise for all method options equally, and many techniques require specialized skills to design, field, and analyze. This limitation can be somewhat mitigated by bringing in expert resources, although programs may also face difficulties gathering that network of expertise.

CUBES and permutations of its methods toolkit have now been used in several large-scale programs, from investigating healthcare provider behavior and household behaviors along the maternal and neonatal healthcare pathway in Uttar Pradesh, India, to understanding and influencing tuberculosis care-seeking in South India (unpublished data) and voluntary medical male circumcision in Africa[5]. We hope that linking the categorization and measurement of enablers and barriers to behavior will enable many more programs to design efficient and effective interventions that get results, and in turn iteratively refine the approaches introduced here.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

Grant information

This work was funded by the Bill and Melinda Gates Foundation and the Surgo Foundation.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The authors thank Prof. Richard Thaler (University of Chicago), Prof. Jessica Cohen (Harvard University), and Anabel Gomez (AVAC) for their insightful comments on the CUBES behavioral framework. We thank Steve Kretschmer, Hannah Kemp, and all past and present Surgo Foundation team members and partners for their contributions to the development of CUBES (formerly called CUE). Finally, we would like to thank Yael Caplan for contributions to the figures.

We thank the Surgo Foundation, who has kindly granted the authors the permission and licence to use materials and information owned by them, for this article.

References

1. Weiss WM, Rahman MH, Solomon R, *et al.*: **Outcomes of polio eradication activities in Uttar Pradesh, India: the Social Mobilization Network (SM Net) and Core Group Polio Project (CGPP).** *BMC Infect Dis.* 2011; **11**: 117.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Semrau KEA, Hirschhorn LR, Marx Delaney M, *et al.*: **Outcomes of a Coaching-Based WHO Safe Childbirth Checklist Program in India.** *N Engl J Med.* 2017; **377**(24): 2313–24.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Thirumurthy H, Masters SH, Rao S, *et al.*: **Effect of providing conditional economic compensation on uptake of voluntary medical male circumcision in Kenya: a randomized clinical trial.** *JAMA.* 2014; **312**(7): 703–11.
PubMed Abstract | Publisher Full Text | Free Full Text
4. Sgaier SK, Reed JB, Sundaram M, *et al.*: **Interventions to Drive Uptake of Voluntary Medical Male Circumcision-A Collection of Impact Evaluation Evidence.** *J Acquir Immune Defic Syndr.* 2016; **72 Suppl 4**: S257–S261.
PubMed Abstract | Publisher Full Text | Free Full Text
5. Sgaier SK, Eletskaya M, Engl E, *et al.*: **A case study for a psychographic-behavioral segmentation approach for targeted demand generation in voluntary medical male circumcision.** *eLife.* 2017; **6**: pii: e25923.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Sgaier SK, Engl E, Kretschmer S: **Time to scale psycho-behavioral segmentation in global development**. Stanford University, and Stanford Center on Philanthropy and Civil Society: Stanford Social Innovation Review. 2018.
Reference Source
7. Bourdieu P: **Outline of a theory of practice**. Cambridge University Press; 1977.
Reference Source
8. Bronfenbrenner U: **Ecological models of human development.** *International Encyclopedia of Education.* 1994; **3**.
Reference Source
9. Rogers E: **Diffusion of innovations**. 5 ed. New York: The Free Press; 2003.
Reference Source
10. Carpenter CJ: **A meta-analysis of the effectiveness of health belief model variables in predicting behavior.** *Health Commun.* 2010; **25**(8): 661–9.
PubMed Abstract | Publisher Full Text
11. Kahneman D, Tversky A: **Prospect theory: an analysis of decision under risk.** *Econometrica.* 1979; **47**(2): 263–92.
Publisher Full Text
12. McCrae RR, John OP: **An introduction to the five-factor model and its applications.** *J Pers.* 1992; **60**(2): 175–215.
PubMed Abstract | Publisher Full Text
13. Prochaska JO, Velicer WF: **The transtheoretical model of health behavior change.** *Am J Health Promot.* 1997; **12**(1): 38–48.
PubMed Abstract | Publisher Full Text
14. Rosenstock IM, Strecher VJ, Becker MH: **Social learning theory and the Health Belief Model.** *Health Educ Q.* 1988; **15**(2): 175–83.
PubMed Abstract | Publisher Full Text
15. Ryan RM, Deci EL: **Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.** *Am Psychol.* 2000; **55**(1): 68–78.
PubMed Abstract | Publisher Full Text
16. Schwarzer R: **Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors.** *Appl Psychol.* 2008; **57**(1): 1–29.
Publisher Full Text
17. Tversky A, Kahneman D: **Judgment under uncertainty: heuristics and biases**. *Utility, Probability, and Human Decision Making: Selected Proceedings of an Interdisciplinary Research Conference*, Rome, 3–6 September, 1973. 1974; 141–62.
Publisher Full Text
18. Tversky A, Kahneman D: **Loss aversion in riskless choice: a reference-dependent model.** *Q J Econ.* 1991; **106**(4): 1039–61.
Publisher Full Text
19. Yzer M: **The integrative model of behavioral prediction as a tool for designing health messages.** *Health communication message design theory and practice.* 2012; 21–40.
Reference Source
20. Dolan P, Hallsworth M, Halpern D, *et al.*: **Influencing behaviour: The mindspace way.** *J Econ Psychol.* 2012; **33**(1): 264–77.
Publisher Full Text
21. Fogg B: **A behavior model for persuasive design.** *Proceedings of the 4th International Conference on Persuasive Technology.* Claremont, California, USA.1541999: ACM; 2009; 1–7.
Publisher Full Text
22. Michie S, van Stralen MM, West R: **The behaviour change wheel: a new method for characterising and designing behaviour change interventions.** *Implement Sci.* 2011; **6**: 42.
PubMed Abstract | Publisher Full Text | Free Full Text
23. Sgaier SK, Baer J, Rutz DC, *et al.*: **Toward a Systematic Approach to Generating Demand for Voluntary Medical Male Circumcision: Insights and Results From Field Studies.** *Glob Health Sci Pract.* 2015; **3**(2): 209–29.
PubMed Abstract | Publisher Full Text | Free Full Text
24. Janz NK, Becker MH: **The Health Belief Model: a decade later.** *Health Educ Q.* 1984; **11**(1): 1–47.
PubMed Abstract | Publisher Full Text
25. India Health Action Trust: **Rolling facility survey: results of two successive rounds, 25 high priority districts, Uttar Pradesh**. Lucknow: IHAT; 2017.
26. Riemsma RP, Pattenden J, Bridle C, *et al.*: **Systematic review of the effectiveness of stage based interventions to promote smoking cessation.** *BMJ.* 2003; **326**(7400): 1175–7.
PubMed Abstract | Publisher Full Text | Free Full Text
27. Spencer L, Adams TB, Malone S, *et al.*: **Applying the transtheoretical model to exercise: a systematic and comprehensive review of the literature.** *Health Promot Pract.* 2006; **7**(4): 428–43.
PubMed Abstract | Publisher Full Text
28. Salmela S, Poskiparta M, Kasila K, *et al.*: **Transtheoretical model-based dietary interventions in primary care: a review of the evidence in diabetes.** *Health Educ Res.* 2008; **24**(2): 237–52.
PubMed Abstract | Publisher Full Text | Free Full Text
29. Kahneman D, Thaler RH: **Anomalies: utility maximization and experienced utility.** *J Econ Perspect.* 2006; **20**(1): 221–34.
Publisher Full Text
30. Neal DT, Wood W, Quinn JM: **Habits—a repeat performance.** *Curr Dir Psychol Sci.* 2006; **15**(4): 198–202.
Publisher Full Text
31. Verplanken B, Orbell S: **Reflections on past behavior: a self-report index of habit strength.** *J Appl Soc Psychol.* 2003; **33**(6): 1313–30.
Publisher Full Text
32. Wood W, Tam L, Witt MG: **Changing circumstances, disrupting habits.** *J Pers Soc Psychol.* 2005; **88**(6): 918–33.
PubMed Abstract | Publisher Full Text
33. Quinn JM, Pascoe A, Wood W, *et al.*: **Can't control yourself? Monitor those bad habits.** *Pers Soc Psychol Bull.* 2010; **36**(4): 499–511.
PubMed Abstract | Publisher Full Text
34. Verplanken B, Wood W: **Interventions to break and create consumer habits.** *J public Policy Mark.* 2006; **25**(1): 90–103.
Publisher Full Text
35. Bronfenbrenner U: **The ecology of human development: experiments by nature and design**. Harvard University Press; 1979.
Reference Source
36. Stokols D: **Establishing and maintaining healthy environments. Toward a social ecology of health promotion.** *Am Psychol.* 1992; **47**(1): 6–22.
PubMed Abstract | Publisher Full Text
37. Reckwitz A: **Toward a theory of social practices: a development in culturalist theorizing.** *Eur J Soc Theory.* 2002; **5**(2): 243–63.
Publisher Full Text

38. Ariely D: **Predictably irrational: the hidden forces that shape our decisions**. 1st ed ed. New York, NY: Harper; 2008; 280.
Reference Source

39. Thaler RH, Sunstein CR: **Nudge: improving decisions about health, wealth, and happiness**. *Rev and expanded ed ed*. New York: Penguin Books; 2009; 312.
Reference Source

40. Brannon LA, Carson KL: **The representativeness heuristic: influence on nurses' decision making.** *Appl Nurs Res.* 2003; **16**(3): 201–4.
PubMed Abstract | **Publisher Full Text**

41. Bränström R, Brandberg Y: **Health risk perception, optimistic bias, and personal satisfaction.** *Am J Health Behav.* 2010; **34**(2): 197–205.
PubMed Abstract | **Publisher Full Text**

42. Laibson D: **Golden eggs and hyperbolic discounting.** *Q J Econ.* 1997; **112**(2): 443–78.
Publisher Full Text

43. Poses RM, Anthony M: **Availability, wishful thinking, and physicians' diagnostic judgments for patients with suspected bacteremia.** *Med Decis Making.* 1991; **11**(3): 159–68.
PubMed Abstract | **Publisher Full Text**

44. Salkeld G, Ryan M, Short L: **The veil of experience: do consumers prefer what they know best?** *Health Econ.* 2000; **9**(3): 267–70.
PubMed Abstract | **Publisher Full Text**

45. Halpern D, King D, Vlaev I, *et al.*: **MINDSPACE: Influencing behaviour through public policy**. Institute for Government. 2010.
Reference Source

46. Ajzen I: **The theory of planned behavior.** *Organ Behav Hum Dec.* 1991; **50**(2): 179–211.
Publisher Full Text

47. Gold GJ: **Review of predicting and changing behavior: the reasoned action approach.** *J Soc Psychol.* 2011; **151**(3): 382–5.
Publisher Full Text

48. Armitage CJ, Conner M: **Efficacy of the Theory of Planned Behaviour: a meta-analytic review.** *Br J Soc Psychol.* 2001; **40**(Pt 4): 471–99.
PubMed Abstract | **Publisher Full Text**

49. Godin G, Kok G: **The theory of planned behavior: a review of its applications to health-related behaviors.** *Am J Health Promot.* 1996; **11**(2): 87–98.
PubMed Abstract | **Publisher Full Text**

50. Brassington GS, Atienza AA, Perczek RE, *et al.*: **Intervention-related cognitive versus social mediators of exercise adherence in the elderly.** *Am J Prev Med.* 2002; **23**(2 Suppl): 80–6.
PubMed Abstract | **Publisher Full Text**

51. Devine-Wright P, Devine-Wright H, Fleming P: **Situational influences upon children's beliefs about global warming and energy.** *Environ Educ Res.* 2004; **10**(4): 493–506.
Publisher Full Text

52. Rimal RN: **Perceived risk and self-efficacy as motivators: understanding individuals' long-term use of health information.** *J Commun.* 2001; **51**(4): 633–54.
Publisher Full Text

53. Bandura A: **Health promotion by social cognitive means.** *Health Educ Behav.* 2004; **31**(2): 143–64.
PubMed Abstract | **Publisher Full Text**

54. McAlister AL, Perry CL, Parcel GS: **How individuals, environments, and health behaviors interact.** *Health Behavior.* 2008; 169–188.
Reference Source

55. Rosenstock IM: **Why people use health services.** *Milbank Mem Fund Q.* 1966; **44**(3:) Suppl: 94–127.
PubMed Abstract | **Publisher Full Text**

56. Harrison JA, Mullen PD, Green LW: **A meta-analysis of studies of the Health Belief Model with adults.** *Health Educ Res.* 1992; **7**(1): 107–16.
PubMed Abstract | **Publisher Full Text**

57. Moreau C, Bouyer J, Goulard H, *et al.*: **The remaining barriers to the use of emergency contraception: perception of pregnancy risk by women undergoing induced abortions.** *Contraception.* 2005; **71**(3): 202–7.
PubMed Abstract | **Publisher Full Text**

58. Curtis VA, Garbrah-Aidoo N, Scott B: **Ethics in public health research: masters of marketing: bringing private sector skills to public health partnerships.** *Am J Public Health.* 2007; **97**(4): 634–41.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

59. Aunger R, Curtis V: **The evo–eco approach to behaviour change**. In: Gibson MA, Lawson DW, editors. *Applied evolutionary anthropology*. New York, NY: Springer New York; 2014; 271–95.
Publisher Full Text

60. Cane J, O'Connor D, Michie S: **Validation of the theoretical domains framework for use in behaviour change and implementation research.** *Implement Sci.* 2012; **7**: 37.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

61. Booth-Kewley S, Vickers RR Jr: **Associations between major domains of personality and health behavior.** *J Pers.* 1994; **62**(3): 281–98.
PubMed Abstract | **Publisher Full Text**

62. Soto CJ, Jackson JJ: **Five-Factor Model of Personality**. In Dunn DS, editor. *Bibliographies in Psychology*. New York, NY: Oxford University Press; 2013.
Publisher Full Text

63. McCrae RR: **The place of the FFM in personality psychology.** *Psychol Inq.* 2010; **21**(1): 57–64.
Publisher Full Text

64. Digman JM: **Personality structure: Emergence of the five-factor model.** *Annu Rev Psychol.* 1990; **41**: 417–40.
Publisher Full Text

65. Følstad A, Kvale K, Halvorsrud R: **Customer journey measures - State of the art research and best practices**. Norwegian Research Council. 2013; Report A 24488.
Reference Source

66. Powell T: **Going beyond surveys for B2B research.** *Quirk's Marketing Research Review.* 2016; **4**.
Reference Source

67. Macdonald EK, Wilson H: **Immersive market research.** Warc Best Practice. 2011.
Reference Source

68. Kawulich BB: **Participant observation as a data collection method.** *Forum Qualitative Sozialforschung.* 2005; **6**(2).
Publisher Full Text

69. Finkler SA, Knickman JR, Hendrickson G, *et al.*: **A comparison of work-sampling and time-and-motion techniques for studies in health services research.** *Health Serv Res.* 1993; **28**(5): 577–97.
PubMed Abstract | **Free Full Text**

70. Manasyan A, Saleem S, Koso-Thomas M, *et al.*: **Assessment of Obstetric and Neonatal Health Services in Developing Country Health Facilities.** *Am J Perinatol.* 2013; **30**(9): 787–94.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

71. Meertens RM, Lion R: **Measuring an individual's tendency to take risks: The risk propensity scale.** *J Appl Soc Psychol.* 2008; **38**(6): 1506–20.
Publisher Full Text

72. Schwarzer R, Bässler J, Kwiatek P, *et al.*: **The assessment of optimistic self-beliefs: comparison of the German, Spanish, and Chinese versions of the general self-efficacy scale.** *Appl Psychol.* 1997; **46**(1): 69–88.
Publisher Full Text

73. Bradley MM, Lang PJ: **Measuring emotion: the Self-Assessment Manikin and the Semantic Differential.** *J Behav Ther Exp Psychiatry.* 1994; **25**(1): 49–59.
PubMed Abstract | **Publisher Full Text**

74. Rimal RN, Juon HS: **Use of the risk perception attitude framework for promoting breast cancer prevention.** *J Appl Soc Psychol.* 2010; **40**(2): 287–310.
Publisher Full Text

75. Goldberg LR, Johnson JA, Eber HW, *et al.*: **The international personality item pool and the future of public-domain personality measures.** *J Res Pers.* 2006; **40**: 84–96.
Publisher Full Text

76. Gregson S, Zhuwau T, Ndlovu J, *et al.*: **Methods to reduce social desirability bias in sex surveys in low-development settings: experience in Zimbabwe.** *Sex Transm Dis.* 2002; **29**(10): 568–75.
PubMed Abstract | **Publisher Full Text**

77. Lowndes CM, Jayachandran AA, Banandur P, *et al.*: **Polling booth surveys: a novel approach for reducing social desirability bias in HIV-related behavioural surveys in resource-poor settings.** *AIDS Behav.* 2012; **16**(4): 1054–62.
PubMed Abstract | **Publisher Full Text**

78. Gregson S, Mushati P, White PJ, *et al.*: **Informal confidential voting interview methods and temporal changes in reported sexual risk behaviour for HIV transmission in sub-Saharan Africa.** *Sex Transm Infect.* 2004; **80 Suppl 2**: ii36–ii42.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

79. Xu T, de Almeida Neto AC, Moles RJ: **A systematic review of simulated-patient methods used in community pharmacy to assess the provision of non-prescription medicines.** *Int J Pharm Pract.* 2012; **20**(5): 307–19.
PubMed Abstract | **Publisher Full Text**

80. Satyanarayana S, Kwan A, Daniels B, *et al.*: **Use of standardised patients to assess antibiotic dispensing for tuberculosis by pharmacies in urban India: a cross-sectional study.** *Lancet Infect Dis.* 2016; **16**(11): 1261–8.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

81. Chan DK, Gallagher TH, Reznick R, *et al.*: **How surgeons disclose medical errors to patients: a study using standardized patients.** *Surgery.* 2005; **138**(5): 851–8.
PubMed Abstract | **Publisher Full Text**

82. Lerman K, Yan X, Wu XZ: **The "Majority Illusion" in Social Networks.** *PLoS One.* 2016; **11**(2): e0147617.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

83. Schiffer E: **Net-Map – Influence mapping of social networks. Manual**. *Sunbelt Conference of the International Network of Social Network Analysis.* 2007.
Reference Source

84. Schiffer E, Hauck J, Andreini M, *et al.*: **Net-map: Influence network mapping–theory**. Small reservoirs toolkit CGIAR challenge program on Water for Food. 2002.
Reference Source

85. Sekandi JN, Zalwango S, Martinez L, et al: **Four Degrees of Separation: Social Contacts and Health Providers Influence the Steps to Final Diagnosis of Active Tuberculosis Patients in Urban Uganda.** *BMC Infect Dis.* 2015; **15**: 361.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

86. Maharana A, Nsoesie EO: **Use of Deep Learning to Examine the Association of the Built Environment With Prevalence of Neighborhood Adult Obesity.** *JAMA Netw Open.* 2018; **1**(4): e181535.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

87. Patodkar VN, Sheikh IR: **Twitter as a corpus for sentiment analysis and opinion mining.** *IJARCCE.* 2016; **5**(12): 320–2.
Publisher Full Text

88. Muzyka CN, Thompson LH, Bombak AE, *et al.*: **A Kenyan newspaper analysis of the limitations of voluntary medical male circumcision and the importance of sustained condom use.** *BMC Public Health.* 2012; **12**: 465.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

89. Quaife M, Terris-Prestholt F, Di Tanna GL, *et al.*: **How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity.** *Eur J Health Econ.* 2018; **19**(8): 1053–1066.
PubMed Abstract | **Publisher Full Text**

90. Quaife M, Eakle R, Cabrera Escobar MA, *et al.*: **Divergent Preferences for HIV Prevention: A Discrete Choice Experiment for Multipurpose HIV Prevention Products in South Africa.** *Med Decis Making.* 2018; **38**(1): 120–33.
PubMed Abstract | **Publisher Full Text**

91. Greenwald AG, McGhee DE, Schwartz JL: **Measuring individual differences in implicit cognition: the implicit association test.** *J Pers Soc Psychol.* 1998; **74**(6): 1464–80.
PubMed Abstract | **Publisher Full Text**

92. Lowes S, Nunn N, Robinson JA, *et al.*: **Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT).** *Am Econ Rev.* 2015; **105**(5): 340–5.
Publisher Full Text

93. Perkins A, Forehand M, Greenwald A, *et al.*: **Measuring the nonconscious**. CP Haugtvedt, P Herr, FR Kardes (2008) *Handbook of consumer psychology.* 2008; 461–475.
Reference Source

94. Sheeran P, Gollwitzer PM, Bargh JA: **Nonconscious processes and health.** *Health Psychol.* 2013; **32**(5): 460–73.
PubMed Abstract | **Publisher Full Text**

95. Oswald FL, Mitchell G, Blanton H, *et al.*: **Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies.** *J Pers Soc Psychol.* 2013; **105**(2): 171–92.
PubMed Abstract | **Publisher Full Text**

96. Friedkin NE, Proskurnikov AV, Tempo R, *et al.*: **Network science on belief system dynamics under logic constraints.** *Science.* 2016; **354**(6310): 321–6.
PubMed Abstract | **Publisher Full Text**

97. Rand W, Rust RT: **Agent-based modeling in marketing: Guidelines for rigor.** *Int J Res Mark.* 2011; **28**(3): 181–93.
Publisher Full Text

98. Pope AJ, Gimblett R: **Linking Bayesian and agent-based models to simulate complex social-ecological systems in semi-arid regions.** *Front Environ Sci.* 2015; **3**: 55.
Publisher Full Text

99. Lee JS, Filatova T, Ligmann-Zielinska A, *et al.*: **The Complexities of Agent-Based Modeling Output Analysis.** *J Artif Soc Soc Simul.* 2015; **18**(4): 4.
Publisher Full Text

100. Fram F, Church F, Sundaram M, *et al.*: **Employing Demand-Based Volumetric Forecasting to Identify Potential for and Roles of Devices in Scale-Up of Medical Male Circumcision in Zambia and Zimbabwe.** *J Acquir Immune Defic Syndr.* 2016; **72 Suppl 1**: S83–9.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

101. Djimeu Wouabe E: **Scoping report on interventions for increasing the demand for voluntary medical male circumcision**. Washington, D.C.: International Initiative for Impact Evaluation; 2013.
Reference Source

102. Population Services International Research Metrics: **Voluntary medical male circumcision TRaC study among men and women aged 15–49 years in Zimbabwe**. Washington, D.C.: PSI Social Marketing Research Series, 2013 2013. Report No.

103. Michie S, Richardson M, Johnston M, *et al.*: **The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions.** *Ann Behav Med.* 2013; **46**(1): 81–95.
PubMed Abstract | **Publisher Full Text**

104. Michie S, West R, Sheals K, *et al.*: **Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used.** *Transl Behav Med.* 2018; **8**(2): 212–24.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

105. Collins LM, Baker TB, Mermelstein RJ, *et al.*: **The multiphase optimization strategy for engineering effective tobacco use interventions.** *Ann Behav Med.* 2011; **41**(2): 208–26.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✔ ? ✔

Version 1

Reviewer Report 29 April 2019

https://doi.org/10.21956/gatesopenres.14025.r27004

✔ **Neela A. Saldanha**
Centre for Social and Behaviour Change (CSBC), Ashoka University, Sonipat, Haryana, India

This is an important and useful article and provides a practical guide to create effective behavior change programs.

The rationale for this toolkit is clearly stated - behavior change interventions often do not work or are not sustainable. We need to move from a "single-intervention" silver bullet approach to a more strategic program approach comprising multiple interventions. But in the face of audience and context complexities, multiple research tools and disciplines, where do we begin? This article provides that answer, both in the form of a practical behavior change framework and appropriate research methods to address each part of the framework.

The behavior change framework the authors outline is extremely useful, as practitioners struggle to cobble together multiple frameworks, none of which comprehensively addresses all barriers. In particular, I appreciated the authors explicitly calling out **contextual barriers & enablers** in the model. Most behavior change frameworks do not reference these factors clearly and yet we know from a body of work on behavioral science that such constraints can be a major driver of behavior and especially of the intent-action gap. Having both perceptual and contextual barriers and enablers in the model makes it a complete framework.

Second, in making the distinction between **habits** or repeated behaviors and one-time behaviors, the authors make the model more clear and effective. Given that habit formation has rarely been a focus in much of development work makes it even more important to explicitly consider whether the behavior we are looking to solve for is a one-time behavior or forming of a habit.

Third, the curated list of research **techniques** the authors outline is of real value to practitioners who need more of these kinds of lists to enable them to expertly navigate the research world. I know I will certainly turn to this list when designing research.

Finally, I must commend the authors for strongly calling for more "in-vitro" **experimentation**. Right now, the only evidence-driven work is through multi-year RCTs that are expensive and not flexible and are only in place after resources have already been expended on designing and piloting the program. Having a way to rapidly test elements of the program at the design stage can ensure we are not wasting valuable resources up-front and can course correct early enough.

As a practitioner, I have a few suggestions to increase the usefulness and likelihood of uptake of the framework by practitioners:

1. **Make it easy:** I would urge the authors to simplify the framework and drive to cognitive ease so that practitioners can easily adopt it. For example, the authors may weigh the benefits of including the "Stages of Change" part of the model against the complexity it introduces for implementation. This could certainly be a more in-depth version of the basic model but in the first instance, defining the behavior well and then understanding perceptual and contextual barriers and enablers to that behavior would go a long way to designing effective programs.

   It would also be helpful to give practitioners some rules of thumb that are actionable. For example, I would recast the narrative of the framework in the form of three simple rules:
   1) Understand determinants of behavior: barriers and enablers, both perceptual and contextual.
   2) Design idea-channel interventions that address barriers and leverage enablers.
   3) Design to all levels of change – individual, family, society.

2. **Consider adding the construct of goals & identity to the perceptual barriers and enablers:** All human behavior is goal driven. Goals drive attention and value and are therefore critical to motivation and establishing intent. At the same time, current research on the field prioritizes knowledge and attitudes rather than goals. In doing so, we often conclude research with little idea of this important determinant of motivation - people's goals and how the intended behaviors can help them achieve their goals. The authors mention the construct of goals several times in the article but should consider including this construct in the framework. Similarly a focus on identity, particularly critical in habit formation would add to the richness of the perceptual barriers and enablers.

3. **Clarify the role of social norms:** It was not intuitively clear to me why social norms which are strongly rooted in psychology are part of contextual factors that has other non-psychological components like laws & infrastructure.

4. **Create more specifics for each construct:** The overall article does a good job of clearly describing each construct. It would be helpful as a job aid to present a specific checklist for each construct. For example, the different types of beliefs one could uncover in research. Practitioners often don’t know all the theoretical underpinnings of a construct and may use their own interpretation. This would result over time in the constructs being interpreted differently and therefore reduce the scope for cross-learning across programs.

5. **Create more detail and specificity on the design process**: The article has a lot of emphasis on the research process to generate insight. There is a much shorter section on effective behavioral design, though this is often the stage when practitioners fail to convert rich insight into effective ideas. Similar to the detailed manner in which the research techniques have been outlined, it would be useful to outline the design process with appropriate checklists. For example how to prioritize ideas, how to effectively prototype prioritized ideas, ways to rapidly test ideas ranging from experiments to rapid qualitative feedback loops and other such steps in the process. Within this design process, there needs to be a clear understanding of the process to design mass media

communication as often mass media & non mass media interventions are created separately rather than as part of one integrated program.

Overall, this is a very practical and clear article. I hope the suggestions I have laid out are useful to the authors and that this work can lead to more such best practices which can help interventionists design effective programs.

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Partly

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Behavior change; behavior science; insights; behavioral design

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 18 Dec 2019

Sema Sgaier, University of Washington, Seattle, Seattle, USA

We thank the reviewer for the positive evaluation of and detailed comments on our work. We appreciate the comments' focus on increasing the likelihood of the framework's uptake by practitioners, and have adapted the article in the following ways:

1. Make it easy: I would urge the authors to simplify the framework and drive to cognitive ease so that practitioners can easily adopt it. For example, the authors may weigh the benefits of including the "Stages of Change" part of the model against the complexity it introduces for implementation. This could certainly be a more in-depth version of the basic model but in the first instance, defining the behavior well and then understanding perceptual and contextual barriers and enablers to that behavior would go a long way to designing effective programs.

It would also be helpful to give practitioners some rules of thumb that are actionable. For

example, I would recast the narrative of the framework in the form of three simple rules:
1) Understand determinants of behavior: barriers and enablers, both perceptual and contextual.
2) Design idea-channel interventions that address barriers and leverage enablers.
3) Design to all levels of change – individual, family, society.

The balance between simplicity and depth is indeed challenging, and we have faced this both internally and when working with practitioners from other organizations. Ultimately, we believe that in this manuscript, there is value in retaining complexity of the drivers at the current level. For example, the present Stages of Change have already been reduced from more nuanced stages in the Transtheoretical Model, and we believe that understanding behavior as a journey from awareness/skills to intention to action and beyond is key to creating successful interventions that highlight that awareness is not enough. However, where we fully agree with this comment is that we need to create better tools to help practitioners engage with the toolkit. Therefore, we have now initiated building a website where we will provide more in-depth and animated walk-throughs of the toolkit, in addition to further case studies.
We also appreciate the actionable three rules of thumb re-framing the narrative of CUBES, and have incorporated them into the Introduction in a slightly modified form:

Ultimately, we encourage practitioners use the toolkit to:
1. *Understand determinants of behavior: barriers and enablers, both perceptual and contextual.*
2. *Design idea-channel interventions that address barriers and leverage enablers.*
3. *Design to all levels of change – individual, family, society, and systems.*

2. Consider adding the construct of goals & identity to the perceptual barriers and enablers: All human behavior is goal driven. Goals drive attention and value and are therefore critical to motivation and establishing intent. At the same time, current research on the field prioritizes knowledge and attitudes rather than goals. In doing so, we often conclude research with little idea of this important determinant of motivation - people's goals and how the intended behaviors can help them achieve their goals. The authors mention the construct of goals several times in the article but should consider including this construct in the framework. Similarly a focus on identity, particularly critical in habit formation would add to the richness of the perceptual barriers and enablers.

These are highly important points. We agree with the importance of seeing behavior as goal-directed, but have not included it as a separate construct, as it folds into the construct of ‘intention’. While goals and intentions are not equivalent concepts, we have now clarified in the Results section and the legend of Figure 1 that intention can mean a plan of action *towards* a specific goal (Ajzen I, Madden TJ. Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. Journal of Experimental Social Psychology. 1986;22: 453–474):
The CUBES framework therefore divides the behavioral change process more simply into three stages of knowledge (ending with the necessary awareness or skills to engage in a behavior), intention (which can also be understood as a plan of action towards a specific goal), and action.

For intervention design, research on specifics of goal-setting has provided fascinating insights: for instance, specific goals are better than vague encouragements to ‘do one’s best’, goals should be challenging yet achievable, and time-proximal goals are more effective than goals far in the future (Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American Psychologist. 2002;57: 705–717). We therefore agree that incorporating the concept of goals into intentions is very useful, although, as we mention below, in this article we cannot focus on the mechanics of intervention design.

We had originally not incorporated the concept of ‘identity’ (such as professional or social identity) into the CUBES framework for two reasons: first, identity can refer to a set of behaviors, personal qualities, or internal beliefs (Cane J, O’Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implementation Science. 2012;7). For intervention design, identity is often thought of more simply as core beliefs about oneself. Second, it is exciting that there are now several studies showing the predictive value of an identity concept on behavior (for example: van der Werff E, Steg L, Keizer K. The value of environmental self-identity: The relationship between biospheric values, environmental self-identity and environmental preferences, intentions and behaviour. Journal of Environmental Psychology. 2013;34: 55–63; Nigbur D, Lyons E, Uzzell D. Attitudes, norms, identity and environmental behaviour: Using an expanded theory of planned behaviour to predict participation in a kerbside recycling programme. British Journal of Social Psychology. 2010;49: 259–284). However, the evidence base on the concept’s predictiveness on behavior is still inconsistent, and too limited in comparison to other types of beliefs to recommend its inclusion into the CUBES framework at this point. We do, however, now mention the emerging evidence in the manuscript (Results):
Beliefs around (professional or social) self-identity may also be predictive of behavior. For example, environmental self-identity, or seeing oneself as ‘a person who acts environmentally-friendly’, is related to several environmental behaviors (van der Werff E, Steg L, Keizer K. The value of environmental self-identity: The relationship between biospheric values, environmental self-identity and environmental preferences, intentions and behaviour. Journal of Environmental Psychology. 2013;34: 55–63). However, empirical research on identity and behavior is still emerging.

3. Clarify the role of social norms: It was not intuitively clear to me why social norms which are strongly rooted in psychology are part of contextual factors that has other non-psychological components like laws & infrastructure.

We thank the reviewer for bringing our attention to this important point. We have now further clarified our rationale in the ‘Results – Contextual drivers of behavior’ section:
Social norms are a construct that can only exist on the level outside the individual: through collective behavior and ‘shared knowledge’, norms describe a set of practices of what other people do (descriptive norms), or prescribe what people should do (prescriptive norms). Both of these may influence attitudes and behavior (Smith JR, Terry DJ, Manstead ASR, Louis WR, Kotterman D, Wolfs J. The attitude–behavior relationship in consumer conduct: the role of norms, past behavior, and self-identity. The Journal of Social Psychology. 2008;148: 311–334). Unlike individual-level beliefs, norms usually imply some consequence to the individual should they deviate from the norm, such as disapproval (Brauer M, Chaurand N. Descriptive norms, prescriptive norms, and social control: an intercultural comparison of people’s reactions to uncivil behaviors. Eur J Soc Psychol. 2010;40: 490–499).

4. Create more specifics for each construct: The overall article does a good job of clearly describing each construct. It would be helpful as a job aid to present a specific checklist for each construct. For example, the different types of beliefs one could uncover in research. Practitioners often don’t know all the theoretical underpinnings of a construct and may use their own interpretation. This would result over time in the constructs being interpreted differently and therefore reduce the scope for cross-learning across

programs.

We fully agree with the reviewer's point that constructs should be defined in greater depth, and indeed they have by the theoretical literature on which this toolkit has been built. We see the most value in building out additional guidance in the form of examples most likely to be useful to practitioners, such as questions that have been or could be asked in surveys to capture constructs. This is a longer-term process, and we aim to incorporate this guidance in a more interactive form on the upcoming web platform.

5. Create more detail and specificity on the design process: The article has a lot of emphasis on the research process to generate insight. There is a much shorter section on effective behavioral design, though this is often the stage when practitioners fail to convert rich insight into effective ideas. Similar to the detailed manner in which the research techniques have been outlined, it would be useful to outline the design process with appropriate checklists. For example how to prioritize ideas, how to effectively prototype prioritized ideas, ways to rapidly test ideas ranging from experiments to rapid qualitative feedback loops and other such steps in the process. Within this design process, there needs to be a clear understanding of the process to design mass media communication as often mass media & non mass media interventions are created separately rather than as part of one integrated program.

In this article, we have aimed to provide an overview of how to systematize potential drivers of behavior and methodologies to generate insights around them. We agree that the ultimate goal is to enable programs to design interventions that address key barriers and leverage main enablers. However, the focus of this article could not be on the specifics around intervention design, for several reasons. One reason is that reviewing the evidence on the intervention types that work or don't in different contexts cannot be done justice in a single article. For example, once a specific social norm has been identified as a barrier to a target behavior, would it work best to leverage key influencers to change the norm, create mass media entertainment showing peer role models (as in an intervention in Rwanda around that used a soap opera to model inter-group harmony; Staub E, Pearlman LA. Reducing intergroup prejudice and conflict: A commentary. Journal of Personality and Social Psychology. 2009), or something else entirely? Behavioral intervention design is its own vast and ongoing area of research and trial and error, and here we can only encourage practitioners to do their own research on potentially effective approaches, once they have used the CUBES toolkit to zero in on the barriers to solve for. However, we have now added some examples of how behavioral drivers have been or could be addressed to the Discussion (see response to the first comment by Reviewer 2). Second, for detail on how to conduct rapid prototyping of solutions, we believe that practitioners of human-centered design, and the frameworks by development partners (such as JHU-CCP, PSI, and FHI360) raised in the second point by Reviewer 1, already provide rich guidance. In contrast, we think this manuscript will be most impactful if focused on categorizing and measuring drivers of behavior.

Competing Interests: No competing interests were disclosed.

Reviewer Report 01 April 2019

https://doi.org/10.21956/gatesopenres.14025.r27007

?

Rashida A. Ferrand

[1] London School of Hygiene and Tropical Medicine, London, UK

[2] Biomedical Research and Training Institute, Harare, Zimbabwe

This paper describes an approach to comprehensively consider drivers of behaviour change to inform effective intervention design.

The authors propose a framework that takes into account both perceptual and contextual factors as well as influencers of behaviour building on available models. This is a refreshing, much needed approach if behavioural interventions are to be effective and lead to impact. I particularly like the emphasis on "repeat and habit" in their framework, in the stages of change.

Issues:

- I am not convinced by some of the approaches (and their use) that the authors present for informing interventions in global health and development. While particular factors may be important, how they can be measured and modified needs further thought. For example, the authors mention use of personality tests and measurement of perceptual barriers; incorporating this in intervention design is not that easy particularly in public health interventions. It is important to consider *how* the findings can be used in intervention design.

- The authors need to consider the limitations of "in vitro" experiments. Behaviour, as the authors point out, is complex with multiple mediators and in vitro experiments may be subject to the same limitations as the approaches used to date, particularly if some factors in the stages of changes are not amenable to change. The issues around effectiveness versus efficacy are still critical in this context.

- The authors present two approaches they have used: 1) in Zimbabwe and Zambia to scale up VMMC and 2) in India on TB care-seeking and maternal and child care. They mention that they have found effectiveness in their own programmes. Disappointingly, there is no evidence presented of the effectiveness of the approach used - the approach in Zimbabwe was published in 2016 and the authors state that it is being piloted nationally in Zimbabwe in 2019. Is there preliminary data on uptake etc.? Available data from DHS etc. in fact suggest that that VMMC rates have fallen in this region. Similarly, the authors present no evidence of the effectiveness of the framework and methods they present. Can the authors demonstrate evidence of the effectiveness of these approaches?

- Limitations about the feasibility of using these approaches and their scientific validity and the feasibility of the findings from these approaches to inform and "modify" the barriers and enablers needs mention and discussion. Ultimately these are the factors that will determine whether the CUBES model is useful.

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology and public health, implementation science, behavioural interventions

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 18 Dec 2019

Sema Sgaier, University of Washington, Seattle, Seattle, USA

We thank the reviewer for the detailed and useful comments on our article, and for highlighting that this approach to approaching behavioral interventions is much-needed. Taking the different points of feedback into account, we have amended the article in the following ways:

1. I am not convinced by some of the approaches (and their use) that the authors present for informing interventions in global health and development. While particular factors may be important, how they can be measured and modified needs further thought. For example, the authors mention use of personality tests and measurement of perceptual barriers; incorporating this in intervention design is not that easy particularly in public health interventions. It is important to consider *how* the findings can be used in intervention design.

We fully agree that measuring and creating interventions around a more comprehensive set of behavioral drivers is not trivial. We think there are two parts to this comment. One is on the research methods themselves. In this article, we have already given nuanced intentions around these approaches: we do not aim to claim that all methods are equally useful, informative, and validated for all global health and development use cases, or even overall. However, we outline a set of options and choices, all of which have been developed, used, and validated in other sectors.

The second part is on then developing interventions that fit the broader spectrum of barriers found, where we agree addressing some barriers may be harder than others, but perhaps the bigger question should be about how *useful* it is to address such barriers. While this manuscript does not address intervention design itself, we mention one approach to categorizing types of interventions by behavioral drivers in the Discussion (Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, *et al*. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Annals of Behavioral Medicine. 2013;46(1):81-95.). In the Discussion, we have now added more detail around what interventions addressing perceptual drivers could look like:

Creating interventions to fit the varied barriers to behavior is a challenge as well as an opportunity for global health. For example, higher levels of conscientiousness have consistently been associated with higher adherence to medication (Molloy GJ, O'Carroll RE, Ferguson E. Conscientiousness and medication adherence: a meta-analysis. Annals of Behavioral Medicine. 2014;47: 92–101) or the contraceptive pill (Leahy D, Treacy K, Molloy GJ. Conscientiousness and adherence to the oral contraceptive pill: a prospective study. Psychology & Health. 2015;30: 1346–1360). Rather than attempting to influence conscientiousness, an intervention might consist of identifying those with lower conscientiousness and targeting increased levels of support to this sub-population. In other situations, drivers may be affected directly. For example, using Facebook to alert college students that peer social norms around drinking were lower than they thought changed drinking behavior (Ridout B, Campbell A. Using Facebook to deliver a social norm intervention to reduce problem drinking at university: social norm intervention using Facebook. Drug Alcohol Rev. 2014;33). As an example of targeting a belief, enhancing self-efficacy through encouragement on progress, attribution of progress to participants' own abilities, observation of others carrying out the target behavior, and other strategies significantly increased physical activity in older adults (Allison MJ, Keller C. Self-efficacy intervention effect on physical activity in older adults. West J Nurs Res. 2004;26: 31–46). These examples are by no means indicative of success in other contexts and behaviors, and interventions all need to be piloted. However, using a framework of behavior allows for the identification of possibilities that may otherwise remain hidden, and conversely narrow the options for choosing suitable intervention types.

2. The authors need to consider the limitations of "in vitro" experiments. Behaviour, as the authors point out, is complex with multiple mediators and in vitro experiments may be subject to the same limitations as the approaches used to date, particularly if some factors in the stages of changes are not amenable to change. The issues around effectiveness versus efficacy are still critical in this context.

We agree with the reviewer that 'in vitro' experiments have inherent limitations, as they are not fully replicative of real-world behaviors in their rich context. We strongly argue that this limitation is different from those of approaches used to date. 'In vitro' experimental approaches provide additional data on which interventions to scale up, as experiments uniquely link controlled inputs to actual behavioral outcomes. Furthermore, a larger set of drivers can be incorporated and tested more rapidly and on a smaller sample than in a field RCT. However, the reviewers raise an excellent point on the validity of 'in vitro' approaches regarding behavior change. We have now added more nuanced, but speculative, guidance on the kinds of behaviors that might especially benefit from 'in vitro' testing to the Discussion:

'In vitro' experiments have inherent limitations, as they are not fully replicative of real-world contexts and behaviors that participants are asked to engage in. However, experiments can link a large set of features to an actual behavioral outcome in a controlled way. It is plausible, but yet to be tested systematically, that the closer the 'in vitro' behavior to its real-world counterpart, the more informative such experiments might be. For example, being asked to pick among different physical products (such as contraceptive packages) with different features is a task not far removed from reaching for an actual product in a pharmacy. Asking community health workers to pick among job options with varying attributes such as salary, workload, and career progression (Abdel-All M, Angell B, Jan S, Howell M, Howard K, Abimbola S, et al. What do community health workers want? Findings of a discrete choice experiment among Accredited Social Health Activists (ASHAs) in India. BMJ Global Health. 2019;4) might not be too different from workers weighing those considerations when looking at a job ad. An experiment asking nurses to evaluate and react to a hypothetical emergency, however, might be much more distant from how the choice behavior would play out in a real-life context.

3. The authors present two approaches they have used: 1) in Zimbabwe and Zambia to scale up VMMC and 2) in India on TB care-seeking and maternal and child care. They mention that they have found effectiveness in their own programmes. Disappointingly, there is no evidence presented of the effectiveness of the approach used - the approach in Zimbabwe was published in 2016 and the authors state that it is being piloted nationally in Zimbabwe in 2019. Is there preliminary data on uptake etc.? Available data from DHS etc. in fact suggest that that VMMC rates have fallen in this region. Similarly, the authors present no evidence of the effectiveness of the framework and methods they present. Can the authors demonstrate evidence of the effectiveness of these approaches?

We appreciate the reviewer's point and fully agree that in this manuscript, the effects, benefits and drawbacks of using the framework have not been discussed in enough detail. This is for two reasons: in the VMMC program, we (the authors) are not involved in the national scale-up and associated data analysis and publication. In our own programs, we and our partners are currently in the process of publishing stand-alone papers where we discuss the approach taken in detail. In particular, in these publications we plan to show how using a structured approach to behavioral drivers enabled data collection and analytic approaches that resulted in finding a) nuanced and actionable predictors of behavioral outcomes, b) distinct segments within the population, and c) a much more specific and segment-tailored space for potential interventions. To caveat this, impact evaluations, i.e. designing these interventions, implementing them, and measuring their effect on behavior or even health outcomes, is a much longer-term endeavor. We are fully transparent that it will take many years and multiple stakeholders to get there.

4. Limitations about the feasibility of using these approaches and their scientific validity and the feasibility of the findings from these approaches to inform and "modify" the barriers and enablers needs mention and discussion. Ultimately these are the factors that will determine whether the CUBES model is useful.

We believe we have outlined, in the brief space available in a single manuscript, the scientific validity (in terms of predicting behavior) behind the behavioral drivers incorporated in the CUBES framework. Detailed nuances, such as sub-categories of drivers, or their usefulness to predict some behaviors more than others, can be found in the vast underlying literature, of which this manuscript can only provide a concise synthesis. It would also go beyond the scope of this manuscript provide an evaluation of each research method in terms of either feasibility or validity. Both concepts are nuanced. For example, feasibility may encompass dimensions such as time, cost, and implementation, design, or analytic skills required. All of these can vary by context, requirements according to research questions and representativeness required, and existing organizational capability. Likewise, the devil is in the detail when it comes to scientific validity of a research method: it would be impossible to say, for example, that quantitative surveys are either scientifically valid or invalid (or anything in between). Instead, the soundness of an item included in a survey will depend on tests of item comprehension, reliability, and validity, and can depend on details such as the polarity of a rating scale used. We therefore encourage readers to take this toolkit as a starting point to judge whether a method might be suitable for their needs, and then use the body of existing literature to go in depth.

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 March 2019

https://doi.org/10.21956/gatesopenres.14025.r27005

Hope Hempstone
United States Agency for International Development (USAID), Washington, DC, USA
Angela Brasington
United States Agency for International Development (USAID), Washington, DC, USA

In general, the article is clear and detailed, making a compelling case for further exploration of the CUBES model. The inclusion of concepts from behavioral economics, in particular, represents a clear departure from other models in common use in global health, and will support ongoing efforts to ensure that behavior change programming reflects current and emerging thinking in behavioral science. Furthermore, the authors' effort to not only describe drivers of behavior, but to present those research methods most appropriate for measuring different drivers, will help render this model actionable and support research-to-practice linkages in behavior change programming. This having been said, the article's rich content - a new theoretical model and a set of under-used tools for measuring behavioral drivers - is worthy of two articles. Each content area has the potential to contribute significantly to behavior change programming; combining them renders them less digestible. We hope that the authors will consider developing a more detailed publication on one or both components moving forward.

We note a number of areas for potential improvement below:

1. While the authors acknowledge the small scale on which this model has been tested in their discussion of limitations, doing so earlier in the article - and discussing how best the model might be further applied within Surgo's work or socialized externally - would help highlight potential for continuing to test and refine CUBES.

2. The authors do themselves a disservice by overstating the novelty of the CUBES model. Many development partners follow a structured process to diagnose and address primary drivers of a

behavior of interest, be they individual, social, or structural. A number of development partners, including organizations like JHU-CCP, PSI, and FHI360, employ models of behavior change similar to CUBES, which synthesize more narrowly focused theories and models. It would be helpful if the authors acknowledged the existence of these models, and articulated how CUBES builds upon and departs from them.

3. Decisions around formative research are driven in many settings by not only theory and methodology, but also by practical concerns regarding cost, speed of implementation, and availability of trained researchers. Acknowledging this reality, and discussing how the proposed approaches might impact these factors, would be helpful to external readers as they consider how best to apply CUBES in the context of their programs.

4. The discussion of social influence could be strengthened (clarified) in a number of key ways. First, the language used implies that influencers are the source of behavioral barriers and enablers ("...these barriers and enablers may be transmitted to the individual through influencers either directly or indirectly or through media channels."). With the exception of behaviors that are strongly (primarily?) influenced by social norms and real or perceived social support, this would seem to be an inaccurate characterization of social influence; would it be clearer and more accurate to note that barriers and enablers are reinforced or contradicted by influencers? The authors also seem to imply that social influence is typically explicit and intentional (as in the context of women's groups, for example), when in fact this is often not the case. Acknowledging the widely varied forms that social influence can take would be helpful to implementers in understanding it as a driver of behavior and considering how best to leverage influencers in interventions.

5. The authors note that one-time and periodic behaviors differ from habitual behaviors, and that different types of interventions may be required to change behaviors depending on their attributes. Attribute-based grouping of behaviors is an area of growing interest among researchers, donors, and behavior change implementers; the authors might consider acknowledging this fact as it has bearing on continued application and refinement of models like CUBES, particularly in the context of multi-health element behavior change programming.

6. The article could be strengthened by more focused attention to the role of emotion in behavior. This driver, in particular, seems to be poorly understood by implementers and researchers; additional guidance on how best to measure emotion would be of tremendous practical utility.

7. The authors and their organization have been vocal advocates for improved audience segmentation in global health programming, and in practice identification of priority behavioral determinants goes hand-in-hand with segmentation of audiences (through doer/non-doer analysis and similar design activities). It would be very helpful if the authors could provide additional detail about the relationship between segmentation and the identification and prioritization of determinants.

8. We beg to differ with the assertion made on p. 13 of the article that randomized control trials predominate in development programming. In our experience, RCTs are employed in only a small minority of programs, particularly in the case of behavior change programming. Recognizing this (and other) practical realities, and addressing them directly, will lend credence to the authors' position.

Is the rationale for developing the new method (or application) clearly explained?

Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
No

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Social and behavior change programming, with particular attention to interventions addressing family planning and reproductive health.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 18 Dec 2019
Sema Sgaier, University of Washington, Seattle, Seattle, USA

We thank the reviewers for taking the time to thoroughly evaluate this article, and for their positive comments. In particular, we are glad the reviewers consider the framework actionable. Taking their feedback into account, we have amended the article in the following ways:

1. While the authors acknowledge the small scale on which this model has been tested in their discussion of limitations, doing so earlier in the article - and discussing how best the model might be further applied within Surgo's work or socialized externally - would help highlight potential for continuing to test and refine CUBES.

We fully agree with the reviewers that continued refinement and testing are crucial to any framework's development. Several levels of testing could conceivably be done, adding different types of value. First, the framework can be assessed on its data collection utility: does it fulfil the promise of speedier data collection development through offering a skeleton of constructs to collect? Through our internal studies (currently in the process of being published), we have found the framework to be helpful in that regard, and have also used it to give rapid feedback on other organizations' research plans.
A second consideration is whether data generated using the framework generates richer insights that increase the intervention options available to programs. In fact, this point was a premise of creating the framework in the first place, as it is plausible that if fewer categories are measured, fewer types of interventions can be matched to them. However, the success of matching distinct

intervention types to distinct enablers and barriers will need to be evaluated through a large set of use cases from many organizations.
Third, there is the question of whether the constructs used in CUBES, taken together, are more predictive of behaviors than those defined in other frameworks. Testing this question is a life-time challenge for researchers and has not been the purview of our work, but we aimed to maximize the likelihood by incorporating only constructs based on a substantial evidence base.
Rather than earlier in the article, we have expanded on these considerations in the Discussion, where we had already touched on them:
The true test of time will lie in whether programs judge CUBES to be useful in increasing intervention fit, as well as effectiveness at reaching target outcomes, as we have found in our own programs. So far, we have used this 'utility test' in our own work, and to give speedy feedback to other organization, in two ways: first, we have used the framework to evaluate planned data collection for comprehensiveness. Second, the framework has been used for dimensionality reduction of expansive datasets (such as household surveys), to enable more clarity in analysis. Future testing should also include critical review of whether intervention options expanded when CUBES was used, and ultimately the change in impact that systematic intervention design yields.

2. The authors do themselves a disservice by overstating the novelty of the CUBES model. Many development partners follow a structured process to diagnose and address primary drivers of a behavior of interest, be they individual, social, or structural. A number of development partners, including organizations like JHU-CCP, PSI, and FHI360, employ models of behavior change similar to CUBES, which synthesize more narrowly focused theories and models. It would be helpful if the authors acknowledged the existence of these models, and articulated how CUBES builds upon and departs from them.

We agree with this important point, and have included the JHU-CCP, PSI, and FHI360 models in the Introduction's overview of behavioral models (however, not in Table 1, as their behavioral driver components are derived from other models):
In global development, several organizations such as PSI, Johns Hopkins' Center for Communication Programs, and FHI360 have created behavioral models incorporating various subsets of drivers. Being application-focused, they usually place great focus on incorporating guidelines on implementation design and monitoring, communication, and advocacy, or on providing a rich compendium of intervention options.

As the reviewers note, these models also incorporate evidence-based drivers of behavior, but are less focused on helping programs systematizing and measuring them. No less importantly, they tend to be more focused than CUBES on either the process of creating change by incorporating guidelines on implementation design and monitoring, communication, and advocacy, or on providing a rich compendium of intervention options. For example, FHI360's 'C-Change' Social and Behavior Change Communication Framework builds mostly on the socio-ecological model and lays out the process of changes management, incorporating some perceptual drivers, but not contextual ones. The Johns Hopkins Center for Communication Programs' (JHU-CCP) Pathways framework, in contrast, starts from laying out different intervention types and what behavioral shifts on individual and community levels can be expected from them. JHU-CCP have also built many successful interventions based on nudges that aim to shift implicit biases, such as commitment devices. PSI's 'Bubbles' Behavior Change Framework is closely related to the COM-B framework, and includes many important drivers, but does not focus on context or influencers and channels.

3. Decisions around formative research are driven in many settings by not only theory and

methodology, but also by practical concerns regarding cost, speed of implementation, and availability of trained researchers. Acknowledging this reality, and discussing how the proposed approaches might impact these factors, would be helpful to external readers as they consider how best to apply CUBES in the context of their programs.

We also agree with the reviewers that decisions about research are impacted by cost, speed, and capability considerations. For this article, we deliberately decided not to go into detail on these points, as the variability in program context, organizational structure, and method chosen all greatly impact these three parameters without many hard-and-fast rules. For example, there cannot be a blanket assessment of whether conducting a quantitative survey or qualitative ethnography are more costly, lengthy, or require additional skills, as it depends on where each program is at a given point in time. Instead, we have added the general point to the Discussion that the three parameters should be considered in tandem at the start of a research process, and that initial resource investment should be compared to later payoffs:
Of course, data collection in any program will also be influenced by considerations about cost, time, and skill resourcing. There is no hard-and-fast rule of how each method ranks on those three parameters, as much depends on existing organizational and program infrastructure. Nevertheless, we urge programs to estimate these parameters before choosing a methodological path, and to also consider trade-offs in investing upfront versus potential time and cost savings in the intervention phase.

4. The discussion of social influence could be strengthened (clarified) in a number of key ways. First, the language used implies that influencers are the source of behavioral barriers and enablers ("...these barriers and enablers may be transmitted to the individual through influencers either directly or indirectly or through media channels."). With the exception of behaviors that are strongly (primarily?) influenced by social norms and real or perceived social support, this would seem to be an inaccurate characterization of social influence; would it be clearer and more accurate to note that barriers and enablers are reinforced or contradicted by influencers? The authors also seem to imply that social influence is typically explicit and intentional (as in the context of women's groups, for example), when in fact this is often not the case. Acknowledging the widely varied forms that social influence can take would be helpful to implementers in understanding it as a driver of behavior and considering how best to leverage influencers in interventions.

The reviewers make an excellent point regarding the clarification of the language around social influence. We agree that influencers can reinforce or contradict enablers and barriers. We think it is not mutually exclusive to say they may also be the source of enablers and barriers: for example, influencers may formulate and transmit certain beliefs. We also agree that social influence is varied and may be explicit and intentional as well as, likely more commonly, implicit. We have now highlighted that variability in the Results section:
Third, these barriers and enablers may be transmitted to the individual, reinforced, or weakened through influencers (such as friends, family, or community members), either directly or through media channels…
Influencers can reach individuals either directly or at scale via various media channels, which is important information to determine how to deliver interventions. For example, female self-help groups can serve as a channel for rural women in India to reinforce or change social norms relating to a certain target behavior. For an intervention to work, the content, the type of influencer, and the channel through which they reach individuals must be identified as relevant to the target

individuals. Finally, social influence is often not intentional as a self-help group might be, but less explicit influence may not be any less powerful.

5. The authors note that one-time and periodic behaviors differ from habitual behaviors, and that different types of interventions may be required to change behaviors depending on their attributes. Attribute-based grouping of behaviors is an area of growing interest among researchers, donors, and behavior change implementers; the authors might consider acknowledging this fact as it has bearing on continued application and refinement of models like CUBES, particularly in the context of multi-health element behavior change programming.

The reviewers' point that attribute-based grouping of behaviors (for example, whether they are one-time or habitual) leads to different types of interventions is well-taken. Intervention design would also be influenced by other types of attributes, such as whether the behavior to be changed is one of commission (doing more of something) omission (refraining from doing something). However, considering the focus of this article – structuring and measuring drivers of behavior – we believe that this would merit a much more detailed exploration in its own article, with a greater focus on intervention design itself.

6. The article could be strengthened by more focused attention to the role of emotion in behavior. This driver, in particular, seems to be poorly understood by implementers and researchers; additional guidance on how best to measure emotion would be of tremendous practical utility.

We fully agree with the reviewers that emotion is a particularly under-measured driver, and that existing methods do not often do a good job capturing it. While this may partly be because of a lack of commitment to capture data about emotion using standard methods such as observation or surveys, the science of measuring emotion is also still very much in development. This is due to the many distinct components and theories of emotion. For example, emotion has an arousal component, which may be picked up using physiological measures such as eye tracking or heart rate. However, arousal does not tell us about valence, or in simple terms whether the arousal is positive (such as excitement) or negative (such as anxiety). Facial expression coding is another possible avenue of emotion measurement, and several private-sector companies now claim to automate emotion labeling from people's facial expressions, for example while they watch a target video. However, there remains a vigorous scientific debate on what these feature detectors actually measure. These methods are also difficult to implement in a field context. For the moment, we believe that adding non-verbal scales, such as those described in the Standardized Scales section, sentiment analysis from speech or text data, as well as simply being conscious of including self-reported emotion in questionnaires, provide the clearest return on investment for programs.

7. The authors and their organization have been vocal advocates for improved audience segmentation in global health programming, and in practice identification of priority behavioral determinants goes hand-in-hand with segmentation of audiences (through doer/non-doer analysis and similar design activities). It would be very helpful if the authors could provide additional detail about the relationship between segmentation and the identification and prioritization of determinants.

We completely agree with the reviewers that identifying priority behavioral determinants and

segmentation often go hand in hand. However, they are not necessarily the same. For example, in some situations it is conceivable that there are barriers to a target behavior that are more or less universal to the population. We usually see segmentation as a next step following from the identification of behavioral drivers within a population: when these data are known on an individual level, segmentation can then reveal sub-groups in the population, and how drivers are differently weighed within them. Behavioral drivers themselves can therefore still be seen as the 'building blocks' of segmentation: for example, one segment of community health workers might have high knowledge and a high self-efficacy being supported by her supervisors and villagers alike, but faces infrastructure problems reaching her village constituents; in another segment, the barriers might be reversed. Our recent methods article on segmentation goes into great detail on how to link the framework of behavior and segmentation, and can be accessed here.
We have added this new reference to the Discussion.

8. We beg to differ with the assertion made on p. 13 of the article that randomized control trials predominate in development programming. In our experience, RCTs are employed in only a small minority of programs, particularly in the case of behavior change programming. Recognizing this (and other) practical realities, and addressing them directly, will lend credence to the authors' position.

While we agree with the reviewers that RCTs do not predominate in practice, we do find that they are seen as the gold standard of evidence, particularly in interventions based on behavioral economics. This may mean programs are missing out on in vitro testing approaches. We have changed the language in the section 'Results - Choosing the right method at the right time for the right purpose' to be more nuanced in that regard:
In global development, many programs tend to focus on descriptive methods for insight generation, followed by field implementation. In the field, randomized controlled trials tend to be seen as the gold standard for assessing the effectiveness of an intervention, even if they are not always employed in practice.

Competing Interests: No competing interests were disclosed.