Dynamics of intervention adoption, implementation, and maintenance inside organizations: the case of an obesity prevention initiative

Mohammad S. Jalali1,§, Hazhir Rahmandad1, Sally Lawrence Bullock2, Seung Hee Lee-Kwan3, Joel Gittelsohn3, Alice Ammerman2

1Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA
2Department of Nutrition, Gillings School of Global Public Health, UNC Center for Health Promotion and Disease Prevention (a CDC Prevention Research Center), University of North Carolina at Chapel Hill, NC, USA
3Department of International Health, Center for Human Nutrition, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

§Corresponding author: MSJ: jalali@mit.edu
30 Memorial Dr, Cambridge, MA 02142

Abstract

Background: Overall impact of public health prevention interventions relies not only on the average efficacy of an intervention, but also on the successful adoption, implementation, and maintenance (AIM) of that intervention. In this study, we aim to understand the dynamics that regulate AIM of organizational level intervention programs.

Keywords: obesity prevention; intervention; sustainability of implementation; motivation; communication; system dynamics; simulation; process evaluation

Methods: We focus on two well-documented obesity prevention interventions, implemented in food carry-outs and stores in low-income urban areas of Baltimore, Maryland, which aimed to improve dietary behavior for adults by providing access to healthier foods and point-of-purchase promotions. Building on data from field observations, seven in-depth interviews, and data discussed in previous publications, as well as the strategy and organizational behaviour literature, we developed a system dynamics model of the key processes of AIM.

Results: We first developed a contextualized map of causal relationships integral to the dynamics of AIM, and then quantified those mechanisms using a generic system dynamics simulation model. With simulation analysis, we show several reinforcing mechanisms that span stakeholder motivation, communications, and implementation quality and costs can turn small changes in the process of AIM into big difference in the overall impact of the intervention. Specifically, small changes in the allocation of resources to communication with stakeholders of intervention could have a nonlinear long-term impact if those additional resources can turn stakeholders into allies of the intervention, reducing the erosion rates and enhancing...
sustainability. Moreover, in the presence of inevitable erosion of intervention components, a minimum level of post-implementation support is required for intervention maintenance.

**Conclusions:** We present how the dynamics surrounding communication, motivation, and erosion can create significant heterogeneity in the overall impact of otherwise similar interventions. Therefore, careful monitoring of how those dynamics unfold, and timely adjustments to keep the intervention on track are critical for successful implementation and maintenance. We discuss theoretical and practical implications.

**Background**
Variation in the successful implementation and maintenance of public health prevention interventions depends, in part, on the organizational dynamics unfolding during and after implementation [1]. The overall goal of this paper is to enhance our understanding of organizational dynamics that impact the effectiveness of obesity prevention interventions, with implications for multi-stakeholder organizational programs in general. Through case-based simulation modelling of complex systems [2], this study contributes to our understanding of the dynamics of program adoption, implementation, and maintenance and thus helps explain why some instances of obesity prevention and treatment interventions prove more effective than others. Specifically, using a simulation model grounded in case data we show how small changes in an intervention can make the difference between failure and success and make the interventions sustainable.

**Obesity prevention interventions**
At its core, obesity is the result of an imbalance between energy intake and energy expenditure in the body, yet multiple factors, from individual biology [3] to the built environment [4-6], social environment [7-9], and economics [10, 11], interact to determine individual energy intake and expenditure. Consistent with this socioecological framework [12], a large number of interventions, from upstream (policy) to downstream (individual level), have been put forward to curb this trend (e.g., see [13-15]). However, the complexity of the obesity problem makes it challenging to design, implement, and sustain successful prevention interventions [16].

Obesity is largely the result of life styles that promote unhealthy eating and limited physical activity within a larger environmental context. Therefore, intervention strategies are moving upstream, with an aim to modify the organizational/community settings in which individuals live, study, and work, in order to limit individuals’ exposure to obesogenic environments and facilitate healthier choices [17-19].

Similar to other public health interventions, many upstream obesity prevention interventions are complex because they share the need to bring together multiple stakeholders to implement new processes and routines within specific organizational and community contexts. According to the RE-AIM Framework [20], the real-world effectiveness of organizational interventions depends not only on the “efficacy” of the intervention in an idealized setting, but also on how well each instance of the intervention (i.e., each program) “reaches” the right population, is “adopted” by the relevant stakeholders, is “implemented” in the organizational setting, and is “maintained” over time to maximize the overall program effectiveness. In fact, a review of the published literature from 1999 to 2010 [21] finds that obesity and physical activity interventions are among the most frequent users of the RE-AIM framework.
Dynamics of adoption, implementation, and maintenance

A key feature of the RE-AIM framework is the shift from short term efficacy (performance under ideal conditions) to longer-term effectiveness (performance under real-world conditions) [22-24]. The efficacy of many lifestyle interventions aimed at obesity and related chronic diseases such as diabetes is well established [25]. Overall effectiveness of these interventions, however, relies not only on the average efficacy of the intervention, but also on the successful adoption, implementation, and maintenance of each instance of that intervention (i.e., each program or organizational level policy change) within the responsible organizational and community context [26]. In practice, much variability in overall effectiveness of interventions arises from variations in AIM. Evidence from a few controlled trials of multiple programs of the same intervention suggests significant variations across programs are common [27, 28].

Figure 1 reflects how multiple instances of the same intervention (i.e., same inherent efficacy) applied to similar population groups (i.e., similar reach) can show different trajectories of overall effectiveness due to their varying levels of AIM.

**Figure 1 Potential program impact trajectories**

For example, one can imagine differences in the implementation process that lead one organization to receive early technical assistance while the other receives none. In the absence of that support, the second organization fails to develop the required knowledge and capabilities, does not receive any positive feedback from its effort, and thus loses momentum. Such mechanisms could lead to significantly diverse results across different instances of AIM for the same intervention. Understanding the sources of variation in AIM is therefore central to enhancing the effectiveness of existing interventions and designing more effective new interventions.

Common effectiveness research methods, such as randomized controlled trials (RCTs), usually focus on measuring efficacy and therefore try to minimize the program variations across AIM by selecting motivated, resource rich, and well-trained organizational participants. Therefore, by design, RCTs may indeed exclude from analysis the very factors that explain
important variations in actual effectiveness of different programs [29-31]. These shortcomings have motivated recent interest in more holistic frameworks to evaluate interventions on dimensions other than efficacy alone (e.g., see [32-36]). Even though there have been several models and frameworks for implementation of obesity prevention interventions, existing research does not explicitly capture the feedback loops between key variables in the system and does not attempt to quantify the key mechanisms of AIM (see Section S1 in the Online Supplementary Materials for more discussion on AIM).

Building on the strategy and organizational behaviour literature

One innovative feature of this study is that it draws on theories in strategy and organizational behaviour to study the effectiveness of chronic disease interventions. Strategy literature indicates that the differences in configurations of organizational resources and capabilities explains much of the heterogeneity in organizational performance [37-39]. Getting to successful configurations of resources is a complex process fraught with many pitfalls, in which some organizations succeed while others fail [40-42]. Building on strategy literature, we look for organizational resources and capabilities which are instrumental in AIM processes of obesity prevention interventions.

A second insight from the organizational literature involves the dynamic trade-offs in building alternative resources. These trade-offs increase the risk of failure in many settings [43]. Organizations are subject to doing what they know best and ignoring new emerging opportunities [44, 45]. They also routinely under-value investments with long-term payoffs [46, 47]. For instance, empirical studies provide strong support for many of quality and process improvement programs [48, 49], yet organizations often fail to fully realize these benefits because resources are withdrawn from the program before full results are observed, initial enthusiasm overwhelms the training capacity to keep the program effective, or seeking short-term gains the system is overloaded with demand and is pushed into a “firefighting” mode of operation [50-53].

Similar to quality and process improvement initiatives, many health care interventions may be beneficial over the long run, but require initial investments and delays before benefits materialize. Therefore, we seek to answer the following questions: What are the mechanisms for building and sustaining the resources central to an intervention’s effectiveness? What are the common failure modes that derail successful development and maintenance of those resources? What are the main leverage points to increase success chances in a program’s life-cycle?

Finally, by explicitly considering the AIM of programs, this study builds on the current public health intervention assessment methods such as RE-AIM [54-56, 20] by using system dynamics modelling to study dynamics of program success and failure. This perspective, combined with model-based experimentation, allows us to develop more holistic theory, evaluate existing and new programs, and provide more operational recommendations.

Dynamic modelling and endogenous perspective

Before discussing our modeling approach, we need to note that our use of the term “model” differs from its common use in public health. In this literature, models are typically event-based, explaining the process of turning research results into practice [57], and are qualitative in nature: Models may lay out sequences of tasks or checklists to facilitate implementation. In contrast, we use models to explicate the causal factors underlying AIM processes (rather than offering a recipe for change), focus on feedback mechanisms that complicate AIM and may lead to policy
resistance [2], and quantify our models to enforce internal consistency and assess the strength of various qualitative mechanisms. Thus, our different purpose for modelling leads us to use the system dynamics modelling approach, which is well-suited for analysing how the effectiveness of organization-level obesity prevention interventions depends on the endogenous dynamics of AIM.

System dynamics modeling and simulation are potential tools to understand the complexity of a system and are increasingly used in public health in general and obesity literature in particular [58-67]. Dynamic simulation models often take an endogenous perspective: they focus on the interactions among concepts within the boundary of the system that lead to behaviors we are interested to understand. This focus does not negate the importance of exogenous drivers of behavior, but is motivated by three considerations [2]: First, the endogenous perspective often provides a richer understanding of the phenomenon, because it does not pass down the explanation to external drivers such as uncontrollable environmental factors or external decision makers. Second, the endogenous perspective brings to the forefront the interactions among various stakeholders that relate to success and failure of AIM, and thus informs modifications that should be sought in response to various signals in the process of implementing and sustaining an intervention. Finally, endogenous dynamics allow us to tease out how otherwise similar organizations can move to different outcomes, a major focus in our study of AIM.

In our setting, the value of simulation is not to produce specific predictions, but to provide a range of likely scenarios and insights into what dynamics drive those scenarios [1]. Dynamic simulation modeling enforces internal consistency in the resulting explanations and allows for quantitative analysis of various tradeoffs. In fact, such models could be used as training micro-worlds [68] to enhance stakeholders’ mutual understanding, commitment, and skills in successfully implementing and maintaining an intervention.

In preparation for discussing the modeling work, we elaborate on two concepts central to system dynamics models: 1) stock and flow variables, and 2) feedback loops. A stock variable represents accumulations and sources of inertia in a system (e.g., a population group, or the number of tasks to be completed in a project). A flow variable represents the rate at which the stock changes (e.g., birth rate is an ‘in-flow’ and death rate is an ‘out-flow’ rate for the stock of population). Feedback loops refer to causal chains of multiple variables in which changes in each variable could be traced back to its historical values. For example, births increase a population which in turn leads to more birth in future, in a reinforcing loop that underlies population growth absent constraints. A feedback loop is called reinforcing if an increase (decrease) in any variable in the loop results in further increases (decreases) when we trace the causal links across the loop back to the original variable. In contrast, in a balancing loop, a perturbation in one variable is attenuated once we trace its impact across the loop back to the original variable. For example, an air conditioner (AC) thermostat creates a balancing loop: An increase in temperature turns on the AC, which brings the temperature down. While balancing loops tend to stabilize systems, reinforcing loops are the source of growth and divergence.

Case study, food environment interventions
The Baltimore Healthy Carry-out (BHC; February-September 2011) and Baltimore Healthy Store (BHS; January-October 2006) interventions were designed and implemented in low-income neighbourhoods of Baltimore (median household income about $20,000 [69]), where carry-out restaurants (a non-franchise fast food outlet) and corner stores are the main food sources [70, 71]. The carry-outs in the BHC intervention are similar to fast-food restaurants but
have different physical layouts—with limited parking availability and no or very few seating arrangements—and lower availability of healthy food choices [69]. They often store, cook, and sell foods behind floor-to-ceiling glass partitions. Many owners may speak English as a second language. Consequently, customers do not interact much with the storeowner or the seller. Storeowners usually know what foods are popular and adjust their menu accordingly. Customers often do not have many healthy options or the choice to request a customized healthier meal [69]. Similar physical layout and language barriers were also observed in the corner stores in the BHS intervention [72, 73].

The scope of the BHC intervention was to design the following strategies: improving menu boards and labelling to promote existing healthier items, introducing and promoting healthy sides and beverages, and introducing healthier combo meals and healthier cooking methods [74]. The aim of the BHS intervention was also to increase the availability and sales of healthier food options in local stores [72] through healthy food incentives (to minimize the financial risk to store healthy items), improving nutrition-related knowledge of storeowners, cultural guidelines to help storeowners build better relationship with their customers regarding cultural and language differences, and guidelines on the types of foods to be stocked and the places where customers could easily access those healthy items [72]. The data for the current study comes from two case studies of BHC and use of archival data on BHS interventions.

Overview of two BHC cases

We studied two cases with different levels of success with implementing the BHC intervention: An African-American food carry-out (a wrap specialty shop) as the successful case and a Korean food carry-out (a fried chicken with sub sandwich shop) as the less successful case. In general, the definition of “success” depends on the subject chosen, and researchers often define various criteria to measure success, such as effectiveness and cost-benefits analysis. Here, we focus on the sustainability of the intervention components as the main measure of success (by comparing the number of remaining components in post-implementation and the total implemented components), while we also acknowledge that other aspects are relevant to defining success.

At baseline, both carry-outs promoted fried food combo meals, offering at least twice as many less healthy foods as healthier foods (see [74] for how food healthiness was defined). They had a limited number of healthy beverages and sides due to a lack of demand or limited storage space. With the help of the interventionist team, both carry-outs initially adopted the major components of the intervention, but later maintained those features at different rates—one kept the majority of the implemented components two years after the end of the intervention, while the other could maintain only about half.

Both BHC and BHS interventions were designed and implemented by researchers from Johns Hopkins University—see (BHC: [74, 75]) and (BHS: [71, 72]) for more information about the design, evaluation, and findings of each intervention. We should note that the BHC and BHS intervention components had a positive impact on healthy food sales and consumer purchases of healthier foods; as a result, the intervention should be considered potentially efficacious (see [75, 72] for more information). Our focus on AIM dynamics allows us to better understand the mechanisms that would impact the ultimate effectiveness of the BHC, BHS, and similar programs.

Methods

Figure 2 summarizes the main processes of our research method. Progress in our method was iterative in nature, such that developing the qualitative and quantitative model exposed new data
needs (hence, the arrows feeding back into earlier steps in Figure 2). These processes are discussed in detail below.

Figure 2 Process flow of our research method. The five processes took place in 2, 2, 1, 3, and 6 months, respectively.

**Data collection through interviews**

We collected data through semi-structured interviews (a common approach for system dynamics modelling [2]) with the key stakeholders in two of the intervention carry-outs in the BHC (after the original intervention implementation had ended)—see the first process in Figure 2. In consultation with original interventionists, we selected two carry-outs with different levels of success with implementing the BHC intervention. Variations across the cases offer useful insights into how AIM processes can diverge.

In our in-depth interviews in the BHC intervention, we interviewed the lead interventionist and two of the carry-out owners in Baltimore that had participated in the intervention. We also interviewed two experts from the Baltimore City Planning Department (the two experts were familiar with the food environment in Baltimore and the BHC intervention). Some stakeholders were interviewed more than once. All interviewees were informed of the purpose and procedures of the research, and assured that the information would be confidential. They signed a consent form and received compensation of $35 per hour of interview (if they accepted the payment) for their time—the study protocol was approved by the Institutional Review Board.

The interviewers used a list of 35 predefined questions and they were free to depart from the list. The list of questions covered various related categories including program history, initial steps in implementation, communication among stakeholders, trust, motivation, post-implementation processes, outputs, financial matters, success and failure instances, and customer feedback (see the questions in the Online Supplementary Materials; Section S4). All interviews were recorded and transcribed for in-depth analysis; the interview with the Korean carry-out owner was conducted in Korean and translated into English.

**Data analysis**

The goal of data analysis was to weave together the case study and published data into a system dynamics model that provides a picture of dynamics relevant to AIM of interventions. Following standards for inductive and generative coding [76], the data were coded by two researchers in MAXQDA 11a to extract: 1) key concepts related to perceived intervention effectiveness; 2) mechanisms of AIM; 3) time-line of events within a program; and 4) quantitative metrics where available. Coding helped in learning the mechanisms of AIM through identifying key variables and relationships among the variables (see the second process in Figure 2). For example, ‘perceived benefits’ (of implemented intervention components) and ‘motivation of stakeholders’...
are two variables extracted from the interviews—which will be discussed in detail—and the relationship between these two variables was that ‘perceived benefits’ had a positive effect on ‘motivation of stakeholders.’ More variables and mechanisms are discussed in the following sections.

In addition to the data collected from BHC, we learned about the intervention from archival data from the original report on the implementation of the intervention (extracted from [77]). We also used published interview data from seven corner stores in the BHS intervention (extracted from [73]). It should be noted that the data from the BHS—an older but similar intervention to BHC (see Section Case study, food environment interventions)—was used as additional supporting evidence to corroborate and contextualize the main case (BHC).

**Qualitative model development**

The synthesis from emerging causal relationships generated a set of dynamic hypotheses, also called a causal loop diagram (see [2] for more information), that provides a qualitative overview of the potentially relevant mechanisms (the third process in Figure 2). These dynamic hypotheses were refined through additional data collection (hence, the feedback from the third to the first process in Figure 2). To this end, we conducted four interviews with the lead BHC interventionist, a support interventionist, and the two Baltimore city experts to get their feedback on the qualitative model draft and elicit further insights. Sample model variables, interview codes, and concrete organizational examples that illustrate the relevant mechanisms are provided in the Online Supplementary Material (S3, Table 1).

**Quantitative model development**

The next step of analysis entailed quantifying these mechanisms into a detailed simulation model (the fourth process in Figure 2). The feedback loops from the previous step summarize the key endogenous mechanisms relevant to understanding the dynamics of AIM in our case study. However, a fully operational simulation model of the AIM dynamics required us to include additional detail in specifying each mechanism quantitatively and include various exogenous drivers, such as the amount of interventionist time available for the intervention—details including model parameters and formulations are fully documented in the Online Supplementary Materials (Sections S5-S9). Operationalizing a conceptual model is the first test; it helps in recognizing vague concepts or areas that were not discussed during the initial data collection and development of the conceptual model. Hence, we reached out to the interventionists to both resolve such ambiguities and parameterize the model.

Quantification of the causal loop diagrams into a simulation model provides a few concrete benefits. First, it enforced internal consistency in conclusions drawn from a complex web of causal pathways, a task the human brain is not well-equipped to do without the assistance of computational tools [78-83]. Second, quantification allowed us to assess the plausibility of various dynamic hypotheses, narrowing down a more complex set of hypotheses to the ones more likely to play a role in actual case histories.

**Model testing, validation, and analysis**

Finally, our quantitative model was tested and validated, and then used for analysis. In general, any model is a simplification of reality and “wrong” in an absolute sense. Thus, the “validity” of a model can only be tested in terms of its usefulness for the purpose of the modelling project. We
conducted various tests to build confidence in the usefulness of our model in understanding the endogenous dynamics of AIM. First, structure and behaviour validity tests were conducted [84]. Mechanisms represented in the model were confirmed in interviews with the stakeholders, including interventionists and experts from the Baltimore City Planning Department. In formulating each equation, we tested it against different input values to ensure it represented the logic portrayed in the data [85], and was robust in extreme conditions. Unit consistency was also enforced in model equations. Extreme condition tests were then conducted to reveal subtle flaws that are not easy to capture by direct inspection or baseline behaviour [2]. These tests specify expected model behaviours under extreme conditions not seen in the field data; e.g., if motivation of stakeholders is zero, no intervention components should be implemented, or if the quality of effort of interventionists is at the maximum possible level and communication among stakeholders is sufficient, there should be no or a few design problems. Behaviour reproduction tests were also used to assess the model’s ability to reproduce the key reference modes observed in various cases and help build further confidence in model’s usefulness [86].

While these tests provided some confidence in the qualitative insights generated from our cases, we should emphasize that the purpose of this study was building theory and not prediction or theory testing. Therefore, in the absence of detailed quantitative data for a larger number of cases, one should be cautious about generalizing the findings or seeking operational advice from our model.

Results
In this section, we present the core mechanisms of the modelb. First, we provide a brief overview, specifying the basic modes of behaviour our modelling work elaborates on. We then discuss the causal feedback loops relevant to understanding AIM that were explicated in the analysis process. Next, we show simulation results that demonstrate the key outcomes of the hypothesized mechanisms and offer insights into the sources of variation in AIM. Additional details including the simulation model consistent with model reporting standards for replication [87] are provided in the Online Supplementary Materials.

Causal feedback mechanisms
In presenting the mechanisms, we draw on the BHC and BHS examples, but provide a more generic terminology and discussion to highlight the transferability of the insights to other interventions that include multiple stakeholders. Below, we discuss the key feedback mechanisms, followed by the simulation analysis. While we present some of the interview data to support our discussions, Section S3 in the Online Supplementary Materials provides more examples of the data, and the process of translating qualitative data into codes, model variables, and mechanisms.

Intervention components, implementation and motivation
An intervention can be seen as a project with a deadline, comprised of various components. Execution of these components, such as designing and installing a new menu board, informs the progress of implementation phase, and depends on the time allocated to implementation by the interventionists, the quality of their effort, and the motivation of carry-out and corner store owners (hereby we call them storeowners) to actively contribute to the intervention. In fact, in the absence of any cooperation by storeowners, no implementation is feasible. In both BHC and BHS interventions, interventionists emphasized building rapport with the storeowners and
Making changes that place minimal burden on the store staff to maintain the motivation of stakeholders when implementing the intervention. Figure 3 summarizes these mechanisms, showing in a box the stock of “Implemented components” that grows with the valve-like flow variable “Implementation rate”.

The first feedback loops in our setting emerge when we consider the impact of implemented components on the carry-out and store operations. Some components may lead to new costs, e.g., for acquiring healthier ingredients and finding new suppliers. Benefits may also ensue, including financial benefits due to increased sales or incentives for participation in the study, reputational benefits, and the personal satisfaction of making a contribution to community health. One carry-out owner in BHC elaborates:

“We could tell from their [customers’] eyes they liked that. My wife is the cashier and some people said that they liked it, and that was encouraging.”

A corner store owner in BHS mentions:

“One day, a new customer just came into the store asking whether we have fruits. I think she just saw the poster (“We Have Fruits” poster) on the door. I also like it. Often my customers gave some positive comments on those.”

An increase in motivation due to observation of such benefits can lead to further implementation of components, and thus even more benefits, in a reinforcing process (Figure 3, reinforcing loop 1 (R1), Seeing the impact). One of the BHS interventionists mentions:

“Even with her limited English, she [storeowner] was trying her best to convince customers to stop at my intervention table on the way out. She would point towards me and the water bottle and try to move them in my direction.”

Strategies such as offering monetary incentives to reduce the financial costs of the program help further increase storeowners’ perceived benefits and eventually increase their motivation in participating in the implementation. One of the corner store owners in BHS elaborates:

“Merchants like us are very sensitive to making a profit. We prioritize our business over everything. So, if you have enough budgets, incentives will motivate stores to participate in the program.”

On the other hand, if the storeowner perceives the costs (both financial and time costs) to exceed those benefits, then a balancing loop may dominate which reduces motivation in response to progress, and slows down further implementation (Figure 3, balancing loop 1 (B1), Costly program). An intervention that, in net, does not benefit the storeowner, has little chance of successful implementation, let alone maintenance.

These mechanisms of perceived benefits are consistent with the discussions presented in a multi-site interview-based study [88] including 52 small urban food stores in eight different cities across the U.S. The diffusion of innovation and implementation science literature also supports these mechanisms. Rogers [89] and Dirksen et al. [90] highlight the idea that cost-effective programs are easier to implement. Legare et al. [91] and Harvey, Kitson [92] also emphasize the perception of intervention as beneficial as a main facilitator for implementation (also, see [93, 94] for more discussions).
Successful implementation depends on the competition between a reinforcing loop (R1: Seeing the impact) and a balancing loop (B1: Costly program). If R1 dominates, successful implementation is possible.

Design quality and communication among stakeholders

The benefits of the intervention, however, are not pre-determined. Design quality is an important aspect of any intervention. A well-designed intervention is less costly to storeowners, may include more benefits, and would be easier to implement and maintain. The quality of design partially depends on the skills and knowledge of the interventionists, which was high in both BHC and BHS cases. Moreover, the intervention should be customized based on the characteristics of each program, and that requires ample communication between the interventionists and the storeowner. Storeowner’s motivation was a major determinant of their availability for communication, and in the BHC case we observed that the owners started with high-levels of motivation. On the other hand, design problems that are not fixed can lead to various issues in implementation and hurt the motivation of stakeholders. Prior research supports the impact of motivation [91, 95-101]. Consider two examples from our cases:

Adding watermelons to the menu during the summer was one of the BHC intervention components and sales data suggested it was profitable [75]; however, it resulted in trash removal problems for the carry-out owner which led to the scrapping of this component: “in the case of the watermelons, during the summer time, even if we wanted to sell those, a lot of garbage would come out of it. Here in [this area of] Baltimore city, there isn’t a place to throw garbage... You can’t put a garbage can outside.”

In the BHS intervention, regardless of the level of storeowners’ motivation and support, they observed some interventionists’ activities as potential barriers to program participation as storeowners were concerned about shoplifting during such activities. One of the corner store owners elaborates:

“It is a big concern how to manage [carefully watch] customers rather than business itself. When interventionists come to the store and talk with customers, it distracts the store business.”

Identifying and addressing such issues requires early communication between interventionists and storeowners, and would enhance design quality. This idea is consistent with prior work highlighting the impact of communication on implementation success [95, 97, 102,
Hence, taken together, communication and design quality create another reinforcing loop (Figure 4, *Communication helps design* (R2)): increased motivation facilitates better communication, which improves design and keeps the stakeholders motivated, in turn keeping storeowners engaged in future communication. Moreover, communication can directly affect motivation through trust-building and sense of engagement (Figure 4, *Motivation through communication* (R3)). Reinforcing loops can amplify small differences between two programs: if one, by chance, faces an early design problem, that can reduce motivation and communication, and sow the seeds of future problems.

**Figure 4** Effect of motivation of stakeholders and communication among stakeholders on design quality. Small problems in design can lead to loss of motivation, reduced communication, and more design problems as implementation progresses.

Note that motivation is also impacted by other factors such as individual knowledge and beliefs about the intervention and self-efficacy to carry out the intervention [104]. While such characteristics have an impact on motivation, we do not explicitly include them in our discussion of endogenous mechanisms because they usually do not dynamically change as part of, and in the time-frame relevant to, the AIM processes within a program.

**Intervention maintenance**

Interventions are not maintained when their components deteriorate, depreciate, or are otherwise scrapped, and are not renewed. For instance, in the BHC, from wear and tear of signs and menus, to changes in prices that may reduce the attractiveness of fresh items, these erosion processes
continually reduce the number of ‘Implemented components’. Yet, the erosion rate is also endogenous, as it depends on motivation, communication, and design problems (Figure 5). Communication can help remind the stakeholders about the need for sustaining changes and fixing emerging problems. High-quality design steps can also foresee, and correct for, the most common modes of failure and thus result in lower baseline erosion rates. Finally, we find that motivated stakeholders are more likely to sustain the changes without external prompts. In the BHS intervention, for example, the storeowners who showed strong or moderate motivation were more likely to help maintain the program (i.e., sustain the stocking of promoted foods) after the program was completed as compared to less motivated storeowners.

These mechanisms create three additional reinforcing loops (Figure 5, R4-R6), as successful implementation raises motivations, improves communication and design, and thus allows for sustaining the gains more effectively and keeping the intervention on track. These mechanisms of the sustainability of the interventions are also supported by the implementation science literature on maintenance—see [95, 93, 105-107] for more discussions.

**Figure 5 Mechanisms affecting the maintenance of interventions.** The erosion rate is influenced by motivation, design quality, and communication, creating three additional reinforcing loops that can drive a wedge between successful and unsuccessful maintenance of programs.
**Simulation analysis**

Given limited space, we only discuss the details of the quantitative model in the Online Supplementary Materials, and here we focus on the simulation results from that model. Our model is generic in the sense that: 1) Its core mechanisms apply to interventions beyond BHC and BHS; and 2) It is not calibrated to any single program, but is parameterized to be qualitatively consistent with the time horizons and magnitudes of variables in the BHC cases. As a result, we focus on understanding the qualitative changes in the model’s behaviour rather than specific numerical values from simulations. This focus thus maintains a balance between the theory-building nature of our paper and the limits of the quantitative data that informs our model.

Our simulation results point to a nonlinear dynamic with potentially important implications for understanding variations in AIM. Specifically, we find that small differences in allocation of interventionist resources to design, implementation, or communication can lead to significant differences in AIM outcomes. We first discuss this tipping dynamic and then present analysis of trade-offs in resource allocations.

**Tipping dynamics**

To demonstrate, consider two identical organizations with identical interventions, composed of various components (e.g., improving menu boards in the BHC program or introducing healthier foods in the BHS program). The only difference between the two simulated organizations is the amount of interventionist time allocated to each (e.g., 8% higher in one case). The project’s adoption, implementation, and maintenance unfold over multiple months: during the first year the focus is on the design of the intervention, the next seven months are mostly focused on implementation, and then the sustainability of the intervention is measured once the interventionists largely leave the scene, offering only some follow up time afterwards.

Figure 6 (A) shows our main outcome variable, the number of intervention components effectively at work in the simulated organizations. This number is zero for the first months as much of the effort goes into designing the intervention. While not shown, the model simulates the quality of designed components in these early months, which impacts implementation and maintenance later on. Implementation starts after about 10 months and speeds up to completely roll out the intervention (i.e., implement its 20 components) by month 19. This period is a burden on the storeowners who invest time and energy without seeing many tangible results, so their perception of benefits goes down early on, only to recover once most of the components are in place. By this time, both simulated organizations show a solid implementation, and if they were actual organizations, they would likely be considered success stories for this intervention. However, what happens afterwards, i.e., maintenance, is key to the long-term effectiveness of each program, and here a small difference in the resources allocated by the interventionist makes a huge difference. One organization keeps most of the components in place, while the other gradually loses most of the components. Why would such a small difference (less than 10% in interventionists’ time during implementation and maintenance phases) have such a major impact?

Early on, the design and implementation processes unfold almost identically for both organizations, and both have enough resources and support to complete the tasks on schedule. The differences become visible only in the maintenance phase. Once implemented, the components are subject to erosion, for example, menu boards may fall and not be replaced and

---

1 The number of components can vary to scale the size of the intervention against resources and other constraints. The number 20 is somewhat arbitrary, but the overall scale is consistent with the timeline of the cases.
healthy items may be dropped from the offering. The rate at which such deterioration happens and the speed with which the required fixes are applied (or ignored) distinguish between organizations that maintain the intervention in the long-run and those that revert back to the old ways of doing things. A few reinforcing loops around communication among stakeholders and motivation of stakeholders are essential for explaining the observed differences in depreciation and re-implementation rates.

**Communication:** First, reinforcing loop “R4-Maintenance reminders” (see Figure 5) highlights the importance of continued communication between interventionists and organizational stakeholders (i.e., storeowners). While the effort needed for keeping this communication is modest, it does provide reminders and support for keeping the erosion rates low, e.g., by fixing any emerging problems before they lead to complete loss of a component. Low erosion, in return, allows for keeping the intervention at its most efficacious state. The intervention thus shows more benefits to the organization, motivates the stakeholders (e.g., storeowners) to keep the communication up with the interventionists, and thus maintains the program in a desired state. On the other hand, a shortfall in communication early after the end of implementation phase can increase the erosion rate, reduce the components standing, cause disillusionment with the program, and further cut down on communication (see Figure 6 (B)).

**Motivation:** A similar mechanism unfolds in reinforcing loop “R6-Left to die” (see Figure 5), as lost motivation increases erosion, reduces the success of the program, and thus further erodes the motivation of organizational stakeholders (see Figure 6 (C)).

**Communication and motivation:** The third reinforcing loop “R3-Motivation through communication” (see Figure 5) connects motivation and communication: a shortfall in communication gradually erodes motivation; that will then require even more communication for fixing task-related issues as well as rebuilding the interpersonal trust and collaborative atmosphere. As a result, the current communication levels fall even further behind what is required, completing a vicious cycle (see Figure 6 (B)). Interestingly, early in the maintenance period the perceived benefits may actually exceed in the ultimately unsuccessful organization (Figure 6 (D)). The reason is that reduced communication early in that period actually reduces the overhead of the intervention for the storeowners, making them potentially more satisfied. Yet, insufficient communication increases the erosion rate which later reduces the success of the program, and ultimately reduces perceived benefits, e.g., due to decreased sales.

As these reinforcing loops take over the dynamics, the unlucky organization falls behind, requires even more time from the interventionist for fixing the problems, which leaves even less time for communication, further strengthening the feedback loops that are affected by the sufficiency of communication. After a few months, the gap between the two otherwise similar organizations becomes very wide and the chances of reviving the intervention in the unsuccessful program remote.

In our simulation experiment, the initial shortfall in communication is triggered by slightly less interventionist time available after implementation is complete (4.8 vs. 5.2 hours per month). However, this small shortfall is amplified through the feedback loops above, leading to the widely different outcomes at the end. On the other hand, for a little while after the completion of implementation phase, the organization with lower interventionist time seems to do even better (see Figure 6 (C-D)), because lower communication translates into less cost for the organization, making the intervention seem more appealing as long as little erosion has happened. The real costs are only revealed once the erosion requires more interventionist time for fixes and thus reduces the sufficiency of communication below acceptable levels.
Note that the exact numbers generated in our simulations are not consequential for the main qualitative finding. While specific numbers vary by the selection of model parameters, the dynamics of AIM under a wide range of parameters include a tipping point that lead to very different outcomes in response to small changes. Lack of attention to the underlying dynamics can lead to erosion of an intervention after it was implemented, in a vicious cycle of lower motivation and communication, faster erosion, and thus less beneficial intervention.

For the intervention to work, a minimum level of communication should be maintained throughout, so that motivation is above a threshold that allows for active support of the implemented components by the storeowner. Such support will then slow down the erosion of the existing components after the active implementation phase is over, thus significantly reducing the ongoing costs of restoring the implemented components. This allows the storeowner and the interventionists to sustain the intervention with limited investment, while keeping up the beneficial impacts, maintaining motivation, and thus locking the system in a fortunate alternative equilibrium.

Figure 6 Implemented components (A), communication sufficiency (B), motivation of stakeholders to implement (C), and perceived benefits (D).
Baseline (blue line) is based on 24 hours effort of interventionists per month. More effort (red line) is based on 26 hours effort of interventionists per month. The big difference between the outputs of these two scenarios relates to the tipping threshold, e.g., a level of interventionists’ efforts that once exceeded causes a sustained intervention. Dmnl: Dimensionless. Y-axis in (A) shows the number of implemented components (capped at 20 components). Y-axis in (B) shows the level of communication sufficiency (>=0), where values below one represent shortfalls in communication. Y-axis in (C) shows motivation of stakeholders (0<= Motivation <=1), where one presents high levels of motivation. Y-axis in (D) shows the ratio of perceived benefits to implementation costs (>=0), where benefits exceed costs for values above one.
Boundary Conditions and Sensitivity of Results

The scenario above highlights the core dynamics of the model using the base case parameter values roughly consistent with the qualitative evidence in the two BHC cases. To gain a more nuanced understanding of the trade-offs involved, we systematically change some of the key model assumptions and assess their impact on the tipping dynamics and the long-term success of simulated programs. Specifically, we explore two questions: What is the impact of interventionists’ quality and capabilities? How do different resource allocation policies influence the sustainability of programs?

Impact of interventionists’ quality and capabilities

The quality of interventionists influences both their productivity in designing and supporting the implementation of intervention components, and the quality of design. In the base case, we simulated very high quality interventionists, but in practice there is often heterogeneity in the capabilities of interventionists, and some are limited in their knowledge, skills, and clarity of communication. As a result, the components designed and seen through by these interventionists may prove problematic, leading to poor implementation or faster erosion. We represent this possibility using a continuous scale changing the quality of interventionist’s efforts \( q \), which scales interventionist productivity and represents the fraction of designed components that are flawless—\( q \) can vary between 0 and 1. We cover the full possible range for this parameter to assess not only case-specific sensitivity of results but also potential mechanisms that may show up in very different cases or extreme conditions. Figure 7 shows that both factors \( q \) and \( p \) (resource allocation in post-implementation) have a significant influence on overall success of the intervention—see Section S2 in the Online Supplementary Materials for more information on how we analysed \( q \) and \( p \) to draw the figure.

Figure 7 Implementation accomplishment in post-implementation. The rate of resource allocation in the post-implementation stage is proportional to that of the implementation stage, such that a value of
50% means that half of resources used during the implementation are allocated in the post-implementation.

Low quality interventionists are not only costly because of their lower productivity, but also because the components they design and implement have many flaws which will in practice reduce the sustainability of implementation. On the other hand, increasing the resources in the maintenance phase would allow the organization to fix the components that erode over time, and keep a higher level of performance in the steady-state, but at increasing costs both to the interventionists and to the storeowners. Moreover, if the interventionist quality is low enough, excess resources cannot solve the problem: the designed intervention will include so many flaws that implementing and maintaining them will require prohibitively large amounts of interventionist time.

Resource allocation policies influence the sustainability of programs
A second way we assess the impact of resource allocation is through considering the impact of giving different priority levels to communication vs. design and implementation of components. In the base run, we assumed communication and design and implementation have similar priorities, so if interventionist time available falls short of what is required for undertaking implementation and keeping satisfactory communication levels, both activities receive the same fraction of their respective desired resources. One may suspect that some interventionists are biased in favour or prioritizing design and implementation, and others may put communication ahead. Would these different allocation strategies offer any advantage or disadvantage? To address this question, we use a more general allocation formulation that changes the priority of communication, \( \alpha \), between 0 and 1. \( \alpha=0 \) represents full priority to design and implementation (so only resources remaining after meeting the demand for those activities are spent on communication) and \( \alpha=1 \) represents full priority to communication, and our base case results are obtained for \( \alpha=0.5 \). Again, we change this parameter between 0 and 1 along with \( q \) (interventionist’s quality), to both observe potential interactions and a comparable graphic (Figure 8).
The contour plot highlights the tipping region we explored in previous section more clearly, where a sudden shift from 90% to 50% sustained components follows very small changes in $q$. Moreover, the impact of priority of communication in resource allocation is nonlinear, and very high and very low priorities lead to poor performance. In fact, the simulated intervention cannot be sustained with single-minded focus on communication because no matter how effective the communication and initial design, the maintenance period entails some tasks that will not be completed in a timely fashion if they have little priority. On the other hand, too little emphasis on communication hurts the initiative by increasing erosion of components due to dwindling motivation, design problems, and lack of communication.

In these simulations the best balance is achieved when communication is given slightly higher priority than implementation and design. Note that the impact of allocation function is modest, compared to the quality of interventionists (or related factors such as the total available resources, or resources available for maintenance). While a better allocation of existing resources, potentially to keep a positive relationship with stakeholders, can help save some marginal initiatives, if a project is plagued with too little resources or low-quality intervention design, resource allocation offers limited leverage.

**Discussion**

This is the first study which: 1) builds on the current public health intervention assessment methods (i.e., through considering the adoption, implementation and maintenance processes) by using system dynamics modelling to study the dynamic mechanisms of intervention success and failure; and 2) draws on theories in strategy and organizational behaviour to study the effectiveness of the interventions.
In this study, based on our in-depth interview data from Baltimore Healthy Carry-outs program and published data from Baltimore Healthy Stores program [73], we developed a system dynamics model showing how the dynamics surrounding communication, motivation, quality of efforts of interventionists, and erosion of interventions can create tipping dynamics which lead to greatly different outcomes over the long haul. Specifically, small changes in allocation of resources to an intervention could have a disproportionate long-term impact if those additional resources can turn stakeholders into allies of the intervention, reducing the erosion rates and enhancing sustainability.

We also found that the quality of the intervention design plays a key role throughout the process. Reviews of research show that health interventions that are designed on a strong theoretical foundation are more effective that those lacking such foundation [108]. Lack of theoretical and practical bases (or poor choice of theory [109]) in the design process leads not only to the ineffectiveness of the innovations in the intervention but also to the ineffectiveness of the implementation process [110, 111]. We showed that a well-designed intervention sustains stakeholder motivation and limits later deterioration; therefore, changes that increase the quality of original design, and maintenance of components, are critical for the long-term success of AIM. Those changes could include the use of more skilled and situationally informed interventionists. They could also include more communication early in the design process with key stakeholders to iterate on the elements of the intervention and to foresee and fix potential problems and gain stakeholder buy in.

While it is easy to call for more resources, in practice most interventions are plagued with budgetary pressures which calls for methods to identify when an intervention is at risk, and how to mitigate those risks. One useful area for improvement is monitoring of stakeholder motivation. This variable plays a key role in the tipping dynamics we identify, if it goes below a threshold, the intervention will become exceedingly costly to maintain. Interventionists should be as sensitive to this motivation level as to the design and implementation of tasks. Intervention design steps may also need to include components that explicitly boost motivation. Again, communication plays an important role in enhancing stakeholder motivation, and thus needs to be prioritized.

Interventionists need to be sensitive to financial or other incentives that stakeholders value, and incorporate them in the design of the intervention to increase the chances that once implemented, the intervention can cross the self-sustaining threshold. For example, in the BHC intervention, certificate from the mayor and the city of the Baltimore was a successful practice that increased motivation at little cost. One of the carry-out owners mentioned: “According to them [interventionists], the certificate was for my contribution to the community with my food. Food that was fresh and good for the community. That made me feel good when I didn’t feel like I wasn’t doing that well. It’s not really about the money to me. I get a lot out of just spending time for a lot of people, finding jobs for people. I like to provide. This is right up my alley. To be recognized for my contribution to the community was really good.”

In addition, finding different mechanisms to motivate stakeholders, especially in the middle of the process (after the early honeymoon period and before they see the actual benefits) can help. These rapport building mechanisms can be frequent site visits by interventionists, asking about ongoing problems and coming up with solutions before tasks are abandoned, and providing data on the benefits (and setting up measurement procedures to track benefits from early on).

A third leverage point is how the design influences intervention erosion rates. Intervention components that can easily become part of the daily routines in an organization (such as the menu board in the BHC intervention which require limited attention for the
maintenance) are much easier to sustain than those that will require conscious and constant attention (such as restocking of baked chips). If organizational routines are to be changed as part of the intervention, structures such as physical layout, supply chains, and decision making processes should be thought through and explicitly designed so that they are consistent with the changes in the core organizational routines. Inconsistencies in those arrangements are likely to increase the speed of erosion of implemented components and diminish motivation over time. For example, if the new process requires new ingredients, the storeowners will find it inconvenient to seek and maintain yet another supplier. The intervention design should follow through with the call for new ingredient to also consolidate the suppliers and avoid future transaction costs to the storeowner (see [112] for an example). The implementation process should also focus on training and empowering organizational stakeholders so that they will appreciate and maintain the components in the absence of the interventionists. Only when the new routines are fully integrated in the organizational culture and processes, can one expect long term sustainment of new interventions.

A common trade-off that our model highlights is the trade-off between designing and implementing intervention components vs. communicating with stakeholders to help build confidence and improve the quality of the intervention. Given that many interventionists are more familiar with the former, there may be a built in bias in the AIM processes against adequate investment in the communication processes central to AIM dynamics. Overcoming that bias and tuning communication levels to address both the motivation and the quality considerations is an important leverage point for training successful interventionists.

Study limitations and future studies
This study is subject to several limitations. On the empirical side, it included post-intervention interviews only. It also included two BHC intervention carry-outs (out of four) as well as only archival data on seven BHS intervention corner stores (out of nine) and did not include control carry-outs or corner-stores. Including control carry-outs could have provided a more nuanced understanding of base-case processes, but was not central to AIM dynamics. In addition to these limitations, the emerging theory we propose can be empirically tested using panel data from a larger number of intervention programs. That would enhance external validity and generalizability. Similarly, the simulation model we developed is simple and generic. More details and numerical calibration to specific cases could be pursued if additional quantitative data was available.

This research can be extended in several directions. Closer integration of quantitative data with the model will enhance confidence in the results. More specific monitoring and performance evaluation metrics could be developed and integrated into protocols for design and implementation of interventions so that corrective action can be taken based on signals observed in each case. Dissemination of a successful intervention design in different organizations may include some interesting dynamics that go beyond the scope of the current paper but are important for overall effectiveness of health interventions. Future case studies could also explore these dynamics in other interventions and organizational settings, expanding the empirical basis for the dynamics, introducing new mechanisms, and offering testing grounds for implications of this research.

Conclusions
Overall, transporting interventions from laboratory settings to community settings is challenging. When the implementation of interventions fails, it is important to know whether the failure
occurred because the intervention was not successfully implemented or if it was ineffective [113]. In many cases, in fact, the intervention is theoretically effective but not properly implemented and maintained. The current study, as a first step towards better understanding endogenous dynamics of organizational interventions, provides evidence on tipping dynamics in health intervention design, implementation and maintenance. The model we develop is stylistic and simple. Real world interventions include many subtle variations and building a fully calibrated model may not be feasible due to data limitations, or may only be viable after the intervention has fully unfolded and the opportunity to improve the situation is lost. Yet our simple model provides a few potential ideas to help monitor and improve the design and implementation of interventions in order to avoid the dynamics that lead to poor long term maintenance of interventions.

**Abbreviations**

**Declarations**

**Ethics approval and consent to participate**
Ethical approval to conduct this study was granted by the Virginia Tech Institutional Review Board (FWA00000572, IRB Number 11-947) in September 2012. The interview data was kept completely confidential: only the researchers with training in protecting human subjects had access to the data, and subjects' information and identity was/will not be revealed to anybody outside of the research team. Informed consent to participate in the study was obtained from all participants.

**Consent for publication**
Not applicable.

**Availability of data and materials**
The simulation modelling files and datasets supporting the conclusions of this article are included within the article and its additional files.

**Competing interests**
The authors declare that they have no competing interests.

**Funding**
Financial Support Provided through National Collaborative on Childhood Obesity Research (NCCOR) Envision’s Comparative Modelling Network (CompMod) program and NIH Office of Behavioural and Social Sciences Research (OBSSR) grant 1R21HL113680-01. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the NCCOR, CompMod, or OBSSR.

**Authors’ contributions**
MSJ collected and analysed the data, developed the model, and drafted the manuscript. HR designed the study, supervised the project and model development, and drafted the manuscript. SLB assisted with the interpretation of the data and editing the manuscript. SHL facilitated data gathering process, contributed her knowledge of the BHC program, and provided critical review of the data interpretation. JG was the PI of the BHC and BHS trials and reviewed and
commented on the manuscript. AA participated in the design of this study and provided insights to enhance the discussions. All authors read, edited, and approved the final manuscript.

Acknowledgements
We would like to thank the three anonymous reviewers for their constructive feedback and suggestions. We are also thankful to participants at the 2014 International Conference of the System Dynamics Society in Delft, Netherlands, the 2014 INFORMS Annual Meeting in San Francisco, CA, the 2014 NIH Conference on Complex Systems, Health Disparities & Population Health in Bethesda, MD, and the 2013 MIT Innovations in Health Care Conference in Cambridge, MA for their feedback on initial reports of this research. We also wish to thank all those who gave their time to participate in this study.

Endnotes
a MAXQDA, software for qualitative data analysis, 1989-2016, VERBI Software – Consult – Sozialforschung GmbH, Berlin, Germany.
b The figures of the model in the article present a general overview of the key loops. For more details of the model, see the Online Supplementary Materials.

References


