Appraising and addressing design and implementation failure in global health: a pragmatic framework

*Ejemai Amaize Eboreime1,2, John Olajide Olawepo3,4, Aduragbemi Banke-Thomas5, Ibukun-Oluwa Omolade Abejirinde6,7, Seye Abimbola8,9

*Corresponding Author: eboreime@ualberta.ca

Authors’ Affiliations:

1. Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada
3. Caritas Nigeria, Abuja, Nigeria
4. Department of Environmental and Occupational Health, School of Public Health, University of Nevada Las Vegas, USA
5. Department of Health Policy, London School of Economics and Political Science, London, UK
6. Centre for Global Child Health, The Hospital for Sick Children (SickKids), Toronto, Canada
7. Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
8. The School of Public Health, The University of Sydney, Sydney, NSW, 2006, Australia
9. The George Institute for Global Health, The University of New South Wales, Sydney, Australia
Abstract

There have been recent concerns about the failure of several global health interventions. Interventions are considered to have failed when they are unable to achieve the intended results. Failure may be linked to how the intervention was designed (design failure) or how it was implemented (implementation failure). Recently, there have been significant efforts to improve the outcomes of interventions. These efforts have led to the development of several theories, models, and frameworks in implementation science to improve the quality of implementation, bridging the divide between evidence and practice. But significant gaps still exist. Whereas much work has been done to develop frameworks and approaches to improve implementation fidelity, not as much effort has been done to guide the adherence of interventions to program theory during the design of the programs. Further, there have been concerns about the applicability of these frameworks in the real-world. This article uses examples to illustrate these gaps and further proposes a pragmatic framework to address identified gaps, thus aiding evidence-informed program design and implementation. The proposed Theory-Design-Implementation (TyDI) framework will support policymakers, program planners and implementers to address potential design and implementation failure, thus improving the fidelity of interventions.

Keywords

Implementation science, fidelity, program planning, health policy, frameworks
Background

Recent years have seen significant efforts to improve the outcomes of global health interventions. The interdependencies between evidence, intervention design and implementation are now well recognized (Bauer et al., 2015; Lahariya & Menabde, 2015). In an ideal situation, the design of interventions (programs and policies) is informed by some form of evidence, which may be empirical or theoretical. Thereafter, these interventions must be deployed and translated by implementers, who are expected to adhere to the design in order to ensure optimal intervention outcomes and impact.

Interventions are considered to have failed when they are unable to achieve the intended results (Klein & Knight, 2005). Failure may be linked to how the intervention was designed (design failure) or how it was implemented (implementation failure) (Allen & Gunderson, 2011). Theories, models, and methods in implementation science to improve the quality of implementation and the effectiveness of interventions have evolved, bridging the divide between evidence and practice (Nilsen, 2015). But significant gaps still exist.

Haynes and colleagues argue that “a strong understanding of the theory of the intervention is a prerequisite for meaningful assessment of implementation, focused not just on the mechanics of delivery, but whether [the] intervention remained consistent with its underlying theory” (Haynes et al., 2016). Thus, a holistic assessment of the success or failure of an intervention will not focus merely on how it is implemented, but also the theoretical basis and fidelity of its design. Whereas some work has been done around the use of systematic approaches like intervention mapping for program design (Fernandez et al., 2019), still much of the scholarship in implementation science has focused on the mechanics of delivery.

Theories, models and frameworks have largely been developed to improve the quality of implementation (implementation fidelity), generating quality evidence from the implementation of
Interventions, and translating evidence to policy makers and implementers (Meyers et al., 2012; Nilsen, 2015; Villalobos Dintrans et al., 2019; Westerlund et al., 2019). But much less attention has been paid to “theoretical fidelity” (aligning intervention design with the theory, logic or hypothesis that informs them), and to the holistic concept of intervention fidelity, which links this theoretical basis with implementation (Gearing et al., 2011; Murphy & Gutman, 2012). This is particularly important for the success of many global health policies and programs (Villalobos Dintrans et al., 2019). This article uses examples to illustrate these gaps and further proposes a framework to address identified gaps, thus aiding evidence-informed program design and implementation.

Identifying core elements of interventions, a dilemma?

Interventions are comprised of “core elements” and “adaptable elements”. Core or essential elements are the components of an intervention that are directly responsible for its impact. Compromising the core elements during the design or implementation phases will most likely result in design or implementation failure, respectively (Fixsen et al., 2009). Adaptable elements of programs include features that make them suitable for the specific implementation context. Adaptable elements can therefore be modified to align with contextual nuances such as local culture, language or socio-political considerations.

Take for example, a recent innovation tested in Nigeria to improve utilization of vaccination – the Vaccine Indicator and Reminder (VIR) ankle band. A core element of the intervention was training mothers on how to activate and read the time strip indicator on the VIR bands. It was however discovered that for cultural reasons, mothers preferred to wear the bands on their children’s wrists, instead of on ankles. This was considered an adaptable element, modifications of which made the VIR bands better accepted by the target population (Obi-Jeff et al., 2020). Thus, evidence ideally informs the core elements of an intervention, while implementation context informs the adaptable elements (Gearing et al., 2011).
Core and adaptable elements may be viewed as answering these questions respectively; “what is being delivered to cause change?”, and “how is it delivered to cause change within context?”. But there are no universal guidelines on how core elements should be identified (Galbraith et al., 2009; Haynes et al., 2016; McKay et al., 2018). This poses a challenge with global health interventions, which are ideally informed by composite social and psychological theories, combining standardised and flexible mechanisms to maximise effectiveness in complex settings (Haynes et al., 2016).

This ‘fidelity versus fit’ problem is a longstanding debate in implementation science. Implementation researchers continue to develop frameworks and models to address the question: “How do we adapt health interventions to fit evolving complex settings without compromising the core elements of the interventions?” (Harn et al., 2013). This is particularly important in global health, given that many policies and programs are commonly designed by actors based in high income countries or in national and sub-national capital cities, but seek to cause change in low- and middle-income countries (LMICs), or at least be perceived as doing so (Rajkotia, 2018).

Limited understanding of context leads to the development of interventions which are often deployed using a “cut and paste” approach (Erikson, 2019); interventions that Olivier de Sardan et al.(2017) referred to as “traveling models”. Local actors, in trying to make these well-funded initiatives work, try to modify the design using “discretionary power” (Lipsky, 2010). In situations where such local actors were not actively involved in intervention design and planning, there is a tendency for only nominal local ownership of the intervention and superficial understanding of its theory of change. This results in a high risk of the core elements being compromised. Therein lies the origin of intervention failures commonly experienced in global health. We illustrate this dilemma further using three real-world case studies.
Case study 1: Primary health care governance reform - Nigeria

Nigeria recently initiated an integrated primary health care (PHC) governance policy commonly known as PHC under one roof (PHCUOR). The policy aims to integrate all primary health care services, structures and subsystems under a single governance body at the subnational level (Odutolu et al., 2016). PHCUOR is premised on a theory of change derived largely from brainstorming by subject matter experts in the national capital city, funded by international donors. Although endorsed by the national council on health for nationwide implementation (National Council for Health, 2011; Odutolu et al., 2016), a recent evaluation revealed that subnational governments not only exercised discretion as to how they implemented the policy, in some cases the policy was redesigned at the subnational level without recourse to the theory of change (Eboreime et al., 2017). This may have untoward implications on the anticipated impact of the policy reform.

Case study 2: Social Health Insurance policy reforms

Increasing interest across many LMICs to improve financial protection and achieve universal health coverage has led to widespread reforms which adopt social health insurance schemes (SHIS) (Obermann et al., 2018; Ogundeji et al., 2019). Ghana, for example, commenced its social health insurance scheme in 2003, funded primarily by taxation through the National Health Insurance Levy (70%), social security contributions (17%), investment income (8%), and premiums and registration fees (5%) (Alhassan et al., 2016; Okoroh et al., 2018). Similarly, Nigeria’s National Health Insurance Scheme was established in 1999, but was officially launched in 2005. The scheme comprises various programs such as social health insurance for formal sector employees, community-based health insurance, private health insurance, and voluntary health insurance (C. Onoka et al., 2014; Onwujeke et al., 2019).

Common across most countries is the inability of these schemes to attain the desired outcomes (Ogundeji et al., 2019). In addition to external actors with varying interests, complexities of health
systems, decentralization, and inadequate capacity of decision makers have been identified as some of the challenges impeding the effectiveness of these reforms (Onoka et al., 2016). This situation is compounded by a lack of clarity by policymakers about the principles of an ideal social health insurance scheme. This has led to policy design flaws responsible for the suboptimal coverage (about 4% of the population) of the scheme in Nigeria (Onoka et al., 2016).

A recent attempt at redesigning the policy in Nigeria is the subnational devolution of Social Health Insurance Scheme. But this devolution comes with a complex fidelity versus fit problem. Nigeria operates a federal system of government with 37 subnational contexts which differ to various extents, socio-culturally and politically (Eboreime et al., 2017). This implies that a centrally designed intervention model cannot be ‘cut and pasted’ across subnational contexts. Rather, redesign of the policy must begin with identifying the core elements of an ideal Social Health Insurance Scheme, and components that can be adapted to various subnational political and sociocultural contexts. Ogundeji and colleagues developed a checklist containing core elements of Social Health Insurance Schemes adapted to the Nigerian context (Ogundeji et al., 2019). The checklist may be used prospectively to guide standardized design and implementation of the reforms, or retrospectively to evaluate theoretical and implementation fidelity.

Case study 3: Disease control programs

Many disease control programs are heavily dependent on donor funding (Das & Horton, 2018) and there are several cases of successful programs. However, several authors have pointed out the lack of available evidence or documentation on failed programs or programs that have not achieved significant impact (Pai, 2019; Rajkotia, 2018). Perhaps, a failure to consider context when proposing new programs may be responsible for such failures (Rajkotia, 2018).
For example, in several projects across LMICs, donors usually gather grant recipients and ‘hand down’ interventions and strategies for project implementation (Abimbola, 2011; Eboereime et al., 2018). In these circumstances, our experience suggests that there is often limited discussion of the theory behind these interventions, and limited understanding of the core and adaptable components of these interventions. Grant recipients may be inclined to move on to implementation without consideration of the core versus adaptable components of these interventions. Adaptations therefore take place along a cascade from the country offices of implementing partners, to the state offices that direct the programs at the subnational levels, then to local offices including health facilities, community- or faith-based organizations. Along this cascade, several changes may occur. The actual intervention may therefore differ greatly from what the donor envisioned. These changes may be additions or subtractions, intentional or accidental (Escoffery et al., 2018), and usually do not have a “common data platform” where they are reported (Chambers & Norton, 2016), or an “adaptation framework” (Escoffery et al., 2018).

The major challenge here is that interventions become adapted several times (Escoffery et al., 2018) and by individuals (on behalf of their own microsystem) of different knowledge gradients, such that the intervention may no longer be recognizable. At the program evaluation (annual or midterm, depending on donor requirements), it is hard to tell if success or failure is because of what was done or in spite of what was done. But such programs usually claim a positive outcome, is as a result of their intervention, or sometimes carry out what Rajkotia (2018) describes as “over-attribute” – i.e. claim for their intervention, a result that is due to another intervention or social process within a country. This ties closely to the discourse on the success cartel in global health and highlights the need for national governments to allocate resources to developing local human capacity to appropriately design, adapt and implement evidence-informed policies and programs (Fonn et al., 2018).
The TyDI framework: Underpinnings, concepts, and parameters

The aforementioned examples demonstrate common challenges faced in global health programming. Clearly, gaps and failures often occur at both design and implementation phases. Many programs and policies are not designed in line with appropriate theories of change, nor with adequate knowledge of the implementation context. Further, an information, communication, engagement or feedback gap is often witnessed between the funders, program designers or policy makers, and the implementers. To bridge this gap, we developed the Theory-Design-Implementation (TyDI) framework which aims to support program planners, policy makers and implementers in the real-world. The framework and its parameters were conceived by the authors through reflection over their experience in the practice of program and policy development and implementation over the years. Each author has about 10 to 15 years of experience in global health policy making and/or program management.

The TyDI framework (see Figure 1), comprises three main elements- a rhomboid (representing real-world program implementation), a hexagon (representing the intervention design) and a circle (the underlying theory of the intervention). Interaction between these elements can be assessed by two indices (implementation index and adaptation index) and two defects (design and implementation).

Figure 1: An illustration of the TyDI concepts and parameters

The framework depicts a design defect; the gap between an intervention as designed (the hexagon) and the theoretical evidence (outer circle) which defines its core components and should inform its theory of change. An ideal intervention aligns optimally with core components of theory both in its design and implementation, but in the real world, design and implementation defects are the norm. The implementation defect, also depicted in the framework, represents the component of the design that was not implemented. This reflects the gap between intervention-as-delivered in comparison to
the intervention-as-designed or planned. The TyDI framework also highlights implementation index, which is a measure of the extent to which the adapted model was implemented in the real world. Likewise, the adaptation index measures the extent to which the designed model was implemented in the real world. Notably, the theoretical fidelity of an intervention can be measured by its adaptation index or design defect, while implementation index and defect can serve as measures of implementation fidelity.

**How to apply the TyDI framework**

We propose that the TyDI framework is best used as a process evaluation framework (Nilsen, 2015). Application of the framework could be prospective or retrospective. Table 1 highlights recommended steps for the application of the TyDI framework.

In view of the fidelity vs fit dilemma, and the absence of universal guidelines to identify core elements of global health policies and programs, syntheses of theoretical and empirical evidence from the literature, in addition to subject matter expert guidance, may be a “best bet” for determining core components of interventions (Haynes et al., 2016). We have not prescribed any specific standards for measuring the TyDI parameters. This is because interventions differ in content and context, thus what works for one may not work for all. The TyDI framework is therefore not premised on a one-size-fits-all paradigm. The aim of the framework is to serve as a pragmatic guide for policy makers and program managers to optimize outcomes and impact of interventions.

**The TyDI Framework in practice**

We applied the TyDI framework to retrospectively evaluate the theoretical and implementation fidelity of a health system performance improvement model in Nigeria called Diagnose-Intervene-Verify-Adjust (DIVA) (Eboreime et al., 2020). DIVA is considered a variant of the Plan-Do-Study-Act (PDSA) cycle, adapted for district health systems (UNICEF, 2012).
First, literature search was conducted to identify the core components of a standard PDSA based on theory. Taylor’s theoretical framework was selected because it was developed from a systematic review of the application of PDSA cycles in healthcare (Taylor et al., 2013). Taylor’s framework proposes five features to test the theory and the application of the PDSA. These features are the use of iterative cycles, prediction-based test of change, small-scale testing, use of data over time, and documentation of processes and outcomes for learning.

Next, the guidelines on DIVA were analysed to identify the design of the model (program theory and internal logic). Further, analyses of program reports, and key informant interviews were conducted to map the implementation of DIVA in Nigeria. Subject matter experts were recruited to determine the theoretical and implementation fidelity of DIVA using a scorecard which evaluated TyDI parameters on a Likert scale. The following operational definitions of TyDI indicators were used to perform the evaluation:

1. Adaptation score/index: This is a measure of the extent to which the steps of DIVA is consistent with the applicable elements of PDSA. It assesses the conceptual similarity between the standard improvement model and its local adaptation.

2. Implementation score/index: This is a measure of the extent to which the adapted model was implemented in the real world. The implementation index may be considered to be a measure of adherence to how DIVA was intended to be used, which is a dimension of implementation fidelity (James Bell Associates, 2009).

Delphi methods (Dalkey & Helmer, 1963) were employed to reach consensus on scores assigned. A team of three subject matter experts was formed and a fourth acted as a process facilitator. The first round of questionnaires was developed using the questions from Taylor’s framework. These questions were tested and adapted to the study context. Each question in the design and implementation was
evaluated on a 3-point ordinal scale (No = 0, Partly = 1, Yes = 2). ‘No’ and ‘Yes’ were attributed to absolute non-compliance and absolute compliance respectively, whereas a ‘partly’ was assigned to partial compliance. The three subject matter experts independently reviewed the qualitative data and assigned a score to each question.

We found that the adaptation indices (theoretical fidelity) of DIVA were optimal (100%) across all core components of the PDSA as defined by Taylor’s theoretical framework. Conversely, implementation fidelity scores were only optimal with two standard features: prediction-based test of change and the use of data over time. Implementation defects of 17% and 50% were found with the other components, the use of multiple iterative cycles and documentation, respectively. Gaps identified in implementation were used as feedback to develop strategies for program improvement.

The scientific validity of the TyDI framework may be of concern to some researchers. Positivists may view the approaches used or suggested for measuring intervention fidelity as not exactly replicable when applied in different settings. However, the approaches proposed or used (such as theory-based evaluations and Delphi methods) are commonly used to address real-world complexities (Birckmayer & Weiss, 2000; C H Weiss, 1999; Carol H. Weiss, 1997). Scientific investigation in a systematic manner does not necessarily inform a ‘good’ theory, rather ‘good’ theory is often discovered through trial and error. This is particularly true with the social sciences and managerial decision-making fields. But ‘good’ theory must indicate how it can be measured for empirical testing (Wacker, 1998). The TyDI framework was designed, not only with the academic community in mind, but as a pragmatic guide to support real world program or policy making and implementation, particularly in global health which is the constituency of the authors. Opportunity for future research exists however, to further test and develop standardized tools and validated measures from the TyDI framework.
Conclusion

How to address persisting inefficiencies and ineffectiveness of global health policies and programs remains a challenge. The search for practical models to improve program or policy outcomes continue in both academic and policy/practice communities. Approaches that can address deficiencies along the evidence to policy and practice continuum may offer solutions to the observed challenges in public health systems. The TyDI framework aims to fill this gap by providing pragmatic approaches to measuring and addressing design and implementation gaps in a way that can be applied by policy makers and program managers in the real world.

References


UNICEF. (2012). *The guidebook: Strengthening district management capacity for planning, implementation and monitoring for results with equity*. UNICEF.


Figure 1: An illustration of the TyDI concepts and parameters

Table 1: Recommended steps to applying the TyDI framework

<table>
<thead>
<tr>
<th>Recommended steps</th>
<th>Some strategies and techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Determine and contextualize the theoretical basis of the intervention.</td>
<td>• Evidence synthesis approaches</td>
</tr>
<tr>
<td></td>
<td>• Gap analysis</td>
</tr>
<tr>
<td></td>
<td>• Participatory stakeholder engagement</td>
</tr>
<tr>
<td></td>
<td>• Intervention mapping</td>
</tr>
<tr>
<td>2. Agree on approach and criteria for measuring TyDI parameters (defects and</td>
<td>• Stakeholder consensus building techniques (e.g. Delphi methods, Nominal group techniques)</td>
</tr>
<tr>
<td>indices)</td>
<td></td>
</tr>
<tr>
<td>3. Determine theoretical fidelity and implementation fidelity from measurements</td>
<td>• Qualitative approaches (e.g. interviews, document analysis, observation)</td>
</tr>
<tr>
<td></td>
<td>• Quantitative approaches (e.g. checklists, surveys)</td>
</tr>
</tbody>
</table>
| 4. Realign intervention to theory and design respectively (prospective or concurrent assessments only) | ● Participatory stakeholder engagement  
 ● Plan-Do-Study-Act cycles |