

Annex A : The Scalability Assessment Tool's terms and concept

The following is an extract of a publication of Management Systems International : Scaling Up : From Vision to Large Scale Change. Tools and Techniques for practitioners, December 2012

Terms

Model : This refers to what is to be scaled up; however, it is not simply the original technical intervention. It is the slimmed down collection of all the inputs and processes both necessary and sufficient to achieve desired results in a variety of settings. Inputs include personnel, funding, and external support (popular and political). Processes include not only technical activities but administrative processes such as human resources, financial systems, and incentives to implement change.

Concepts

1. Based on statistically significant, sound evidence of sizable impact. This refers to scientific evidence. The more credible the evidence, the easier it is to convince policy and decision makers and other stakeholders of the value of the model. Statistically significant quantitative or qualitative evidence of impact is preferable. It is important to note that while impact data is more desirable than outcomes, and outcomes more desirable than outputs, etc., it is often the case that while pilots are constrained in proving impact so that "lower" levels of evidence are accepted, it is only acceptable if the logic connecting lower (outcomes) and higher (impact) levels is clear and confirmed by other studies or international evidence.

2. Independent evaluation by respected and reliable sources. Evidence from independent sources, such as an independent external evaluation, increases credibility. Often pilots rely on data conducted, collected, or analyzed by the organization itself, which even when of high quality and good design, can be perceived as biased. For scaling up purposes, there is a strong preference for a formal external evaluation of the model. A second best alternative is the rigorous external review of internal evaluations combined with an external advisory panel in the design and implementation of the research design and monitoring and evaluation framework. Publication in a well-regarded peer-reviewed journal can usually be considered equivalent to an external evaluation or review.

3. Model works in diverse social contexts. A model works in diverse social contexts because it has been successfully implemented in multiple and diverse settings, or within a wider institutional context, such as access to and quality of services or commodities. Examples of relevant social variables might include population density, attitudes on family planning and size, degree of poverty, and access to services, etc. Examples of relevant institutional variables are: availability of transportation, availability and cost of family planning commodities, and quality of care. For the purposes of scaling up, there is a strong preference for projects that have been piloted in multiple sites and settings, i.e., that have both internal and external validity. This is particularly true in countries like India or Nigeria where even within national boundaries there is huge diversity in social norms, infrastructure, governance, incomes, etc. The following list combines measures of internal and external validity. The further up the ladder of success listed below, the stronger the candidate.

- ❖ Innovation (minimal objective evidence)
- ❖ Promising Practice (anecdotal reports)

- ❖ Model (positive evidence in a few cases)
- ❖ Good Practice (clear evidence from several cases)
- ❖ Best Practice (evidence of impact from multiple settings and meta-analyses)
- ❖ Policy Principle (proven, a “truism” essential for success)

4. The model is supported by eminent or credible individuals and institutions. Support or endorsements from eminent individuals or institutions help in advocacy efforts. These can be either recognized experts who confer legitimacy, e.g., the Indian College of Medical Research or the World Health Organization, or well-known and famous personalities who confer status, whether from the media, sports, politics, or other fields, e.g., first ladies, or the Indian actor/model Amitabh Bacchan’s support for polio vaccine.

5. The results are visible to casual observation, tangible. It is easier to convince people of the impact of things that they can easily see and experience for themselves than those they cannot. Experience has shown that it is easier to scale up interventions like bednets or treatments that stop maternal hemorrhage than those where the results are less visible or tangible, such as the empowerment of women or communities.

6. Results are clearly associated with the intervention. It is easier to convince people if the impact is not just a byproduct of an intervention, but easily attributable to the model itself. This criterion emphasizes the fact that decisions to adopt and implement models are often not made, or are only made, on the basis of objective, statistical evidence despite an increased emphasis on high-quality evidence and evidence-based policy making. Outcomes in health, education, or other fields are often driven by multiple factors. Skeptics may claim that the results observed were not due to the intervention but to other factors. While a good experimental or quasi-experimental design can address this for a technical audience, the ability to literally see the causal relationship can be an important factor for a non-technical audience. Models that are difficult to scale up are those in which impact is lagging or delayed or the causal mechanisms are not straightforward, e.g., the effects of psychosocial development on childhood performance. Counterexamples are Oral Rehydration Therapy (salts) to reduce childhood diarrhea, an intervention with a direct, immediate, and simple causality. Sometimes models are able to go to scale, or at least convince key decision makers to approve adoption and funding, without strong evidence because their internal logic is inherently compelling; this is often true of models that involve training or have evidence from the literature. However it is often the case that there are several other assumptions which need to hold true for the causal chain to achieve desired outcomes, and it is important to test the validity of these assumptions.

7. Evidence and documentation that have a strong emotional appeal exist. This criterion emphasizes the fact that decisions to adopt and implement models are often not made on the basis of objective, statistical evidence but on emotions (despite the increased emphasis on high-quality evidence and evidence-based policy making). Seeing a sick baby healed or a woman who is bleeding to death revive and recover could be more important than dry statistical evidence to many policy and decision makers. Successful advocacy uses both scientific and other evidence to advance change, especially through stories, photos, video/film, and other media that are able to convey a message with an emotional appeal.

8. Addresses an objectively significant, persistent issue. It is easier to successfully advocate for social problems and issues when those issues affect large numbers of people in a significant way. In other words, it is easier to scale up models that an objective outside observer would rank as one of the major challenges in health, education, etc., for a country, population, or subregion. In health, for example,

this may be an issue that is a leading cause of death or recognized as an easily preventable or treatable cause of death or disease with permanent debilitating effects. That is why tuberculosis, malaria, and polio have all become major health priorities. In education, this may be levels of malnutrition and stunting in young children with the effect of permanent and destabilizing mental abilities. It is important to point out that just because an issue is objectively important, it does not mean that it is a major policy priority. One of the actions that can come out of a scalability assessment is precisely what needs to be done to move an objectively important issue higher up on the policy and social priority list. Early childhood education is a good example of an objectively important issue which, in many cases, has not received adequate attention.

9. Addresses an issue that is currently a high (policy) priority of potential adopters and is aligned with organizational goals, mission, and vision. This item is particularly relevant when scaling up is being considered by the government, where government approval, funding, or other roles may be important. When an issue is high on the policy agenda, it means that the government is actively looking for solutions and willing to provide funding. It is easier to scale up something which is already an important priority than to move something up the agenda. For example, a current environmental intervention that addresses maternal mortality is also a Millennium Development Goal; therefore, it is easier to scale up than interventions that address the health of gay, lesbian, bisexual, and transgender populations. It is a plus when organizations and their constituencies already consider the issue a high priority and when it is already aligned with the organization's mission and vision; thus, expansion is facilitated when an organization's funders or board of directors think the issue is important.

10. Addresses a need sharply felt by potential beneficiaries or participants in the target area. A felt need means that if survey research, focus groups, or other forms of interviews were conducted, the issue that the model addresses would emerge from participants as a major priority. It is easier to mobilize popular and grassroots support for interventions that address felt needs to achieve acceptance of new programs and interventions. Felt needs tend to be tangible and immediate like poverty and livelihoods, safe drinking water, and curative care as opposed to issues where benefits are captured in the future, e.g., preventive care, or where there lacks the perception that the status quo could or should be different, e.g., levels of infant or maternal mortality, or number of births per family.

11. Current solutions for this issue are considered inadequate. Current solutions refer to programs or policies already in place to address the issue. It is easier to advocate for the model if people and decision makers assess that the current solution is not working, or if there are no solutions at all. If there are no current solutions in place, this concept refers to whether people think the lack of a solution is a problem at all, e.g., providing child health interventions solely through facility based approaches.

12. Superior effectiveness to current solutions is clearly established. This criterion should be self-evident. If the model being proposed for scale up has evidence of greater impact or effectiveness than existing solutions, it is easier to scale up.

13. Superior effectiveness to other innovative models established. Advocacy of a particular model or solution must compete with other approaches for the attention of policymakers. The model should be able to demonstrate success in achieving results, but also superiority over alternative approaches. It is also important to assess if there are any foreseeable potential opposition from vested interests or social/cultural groups. Some issues or solutions (the intervention) can threaten vested interests or

become controversial for certain elements of the population. In public health, interventions which shift tasks from skilled to less-skilled professionals can threaten the former's authority, prestige, and perhaps most importantly, livelihoods, even when they are already overburdened. For example, having VHWs in rural areas give injections of antibiotics to newborns with sepsis was opposed to by neonatologists, even though there was no evidence of complications in the administering of injections nor were there any neonatologists who actually served those populations. Similarly, (unqualified) rural medical practitioners have opposed the management of public primary health clinics by NGOs because they fear the loss of business if the quality of care, and therefore competition, improves. In many areas, adolescent sexuality and health education can be opposed by conservative social or religious groups on ethical or religious grounds. The lack of opposition can increase the prospects for scaling up. It is important to establish some clear standardized criteria for comparison with a common denominator, preferably in terms of outcomes and impacts.

14. Implementable within existing organizational systems, infrastructure, and human resources. It is easier to implement a model at scale if it can use existing systems, i.e., if they can be grafted on, then if they require the creation of a whole new infrastructure, organization, management, and systems. This is because investing in new systems and infrastructure is usually expensive and can be politically threatening to existing bureaucracies and organizations. However, on some occasions, creating a new system may have advantages in terms of circumventing political and bureaucratic obstacles and weak capabilities. Models which require long and intensive training periods for staff are more difficult to scale up and face opposition from potential adopters due to the costs and time involved.

15. Contains a few components easily added onto existing systems. The fewer the elements of a model, the simpler it is to scale up. The Home-Based Newborn Care model that was originally implemented had several distinct components which made large-scale implementers wary of scaling it up. Thus, while designing a scaling up strategy, it was important to simplify the model so that implementation on a larger scale was easier.

16. Small departure from current practices and behaviors of the target population. It is easier to implement a model if does not require significant changes from the existing behaviors, beliefs, and practices of the target population, clientele, or beneficiaries of the model, e.g., counseling mothers on using food cooked for the family as supplementary feeding for children, versus advocating the cooking of meals separately.

17. Small departure from current practices and culture of the adopting organization(s). This concerns the intervention's compatibility with the culture, practices, physical infrastructure, and resources of the adopting organization(s). It is easier to implement a model if does not require significant changes from the existing behaviors, beliefs, and practices of the organization that will implement the model at scale. For example, in a government system that emphasizes delivery of curative medical services, the introduction of community mobilization, treating clients with respect and dignity, preventive services, or working with and through non-professional VHWs may require a substantial changes in culture or norms. It is important to note that even in the case of Expansion, where the implementing organization remains the same, a major challenge may be in maintaining its organizational culture, practices, and roles as the organization grows. This criterion also assesses the extent to which adoption/scaling of the model is likely to meet internal opposition, resistance, or support within the adopting or expanding organization.

18. Few decision makers are involved in agreeing to the adoption of the model. This criterion addresses the number and hierarchy of decision makers required to approve scaling up. For example, the Home-Based Newborn Care model required not just the approval and buy-in of policy makers and programme implementers in the public health system, it also needed the approval of the Indian Academy of Paediatrics and the National Neonatology Forum. This process of engagement and negotiation with multiple decision makers can be tedious and slow the scaling up process.

19. Demonstrated effectiveness in diverse organizational settings. Compatibility with systems and infrastructure within diverse organizational settings is best demonstrated by actual evidence in those types of settings. This can help convince organizations and skeptics in general that the intervention will work, and serves as a strong counterweight to the “not invented here” argument—indicating that there is evidence that the model has had similar impact when implemented in multiple and diverse organizational settings whether it be a NGO, government, or private sector actor; or a tertiary, secondary, or primary facility in health. Diversity can refer to variances in quality in terms of infrastructure, equipment, supplies, and the effectiveness of the organization that is implementing the model. The more evidence available on the robustness and resilience of the model in different organizational settings, the easier it is to scale up.

20. The model is not particularly value or process intensive. Here, value or process are defined as models that focus on tacit knowledge, i.e., how things are done or delivered versus the technical content (which focuses on drugs, equipment, procedures, and information such as changing attitudes, knowledge, and practices). Value or process intensive models are much more difficult to implement at large-scale because: (a) quality is usually important, (b) they tend to be time and resource-intensive to transfer to others, and (c) are often a substantial departure from existing practices in organizations with large-scale capacity because they are difficult to systematize, relying heavily on organizational culture. Examples of values or process intensive components include: community participation, community mobilization, teaching service providers to treat beneficiaries with dignity or respect, and free play in early childhood education versus lectures and rote learning.

21. Low technical sophistication of the components and activities of the model. Models can be sophisticated or not whether they are technology/content-intensive or process intensive. The more sophisticated they are, the more they require either higher skilled and trained implementers or greater investments in the training and capacity building of personnel to implement them successfully. Thus, less sophisticated models are easier to scale. For example, standardized doses of drugs that require only one dose, do not require cold chains, are the same dosage regardless of body weight, and do not need monitoring of potential negative after effects are easier to scale up than those where multiple elements require a number of complex decision points.

22. Key innovation is a clear and easily replicated technology. A model which has a single element such as delivering a vaccine through an existing service delivery mechanism stands a much better chance of approval for scaling up. For instance, a childhood pentavalent vaccine was introduced only in states that had robust immunization programs, where it was perfected over several decades and easily scaled up with minimal modifications to the infrastructure and reporting mechanisms.

23. Low complexity model, simple with few components added onto existing systems. Models that are complex, have multiple elements, and need updated delivery mechanisms often require long lead-in times during which permission is obtained from gatekeepers and leaders, and acceptance and buy-in has to be created among participants, beneficiaries, and community leaders. In some cases, this can

take a year or more; in other cases, it is not necessary because piloting organizations have already established trust and credibility based on their long-term relationships and presence in the localities in which they work. Regardless, models which require the building of trust, relationships, and permission from gatekeepers are harder to scale up or at least, take much longer. Models vary widely in terms not only of the pre-conditions necessary before a model is implemented, but the training, skills building, and general capacity required of front-line implementers, supervisors, and other staff. Capacity building is time, effort, and resource-intensive and often difficult to replicate at scale when all three of these elements are scarce.

24. Model requires little supervision or monitoring. Models which require intensive monitoring and supervision are harder to scale up. Many organizations with large capacity are weak in these areas, and it can be difficult to enforce monitoring and supervision at large-scale. Models in which implementers are either self-motivated or where they can work effectively independently are easier to scale up. Monitoring intensive models can be made simpler by the use of monetary and non-monetary incentives, and by using electronic data collection systems supported by on-site supervision as required.

25. Able to be tested by users on a limited scale. This means that organizations that are interested in adopting the model should be able to run their own pilots without having to commit substantial resources or undergo significant organizational changes to see if it works.

26. Superior cost-effectiveness to existing and competing solutions clearly established. Models not only need to be more effective than existing or competing solutions to have a chance of being scaled up, they need to be more cost-effective. This requires getting comparable measures of cost-effectiveness, which can be hard to come by in low-resource situations.

27. Requires a large commitment of funds at scale. Models that require substantial upfront commitments are generally complex and require changes in existing systems. Advocacy for this is facilitated by demonstrating high impact, extensive stakeholder support, and importantly, a good understanding of the unit cost (or some other measure of impact) per resource spent. This needs to go beyond simple calculations such as the total project budget divided by the number of beneficiaries, communities, etc., which often underestimates costs by ignoring the contributions of management and technical assistance or which overestimates them by including evaluation costs which will not be replicated at scale. Having solid cost data is necessary for successful advocacy, and knowing whether or not an intervention is affordable given the available resources and objective need can make or break an intervention in terms of scalability.

28. The model has its own internal funding (e.g., user fees), corpus or endowment, or some other long-term sustainable funding source, or overall funding for this sector is a priority. Funding here refers to the funding of the model and not the project which piloted the model. It means that the model includes an internal mechanism of recovering costs whether they are user fees, profit margins on goods or services provided, etc. Overall funding for this sector is a priority and growing, to the extent that there are unspent funds. Funding refers to public sector support for the general sector in which the model works, e.g., health, education, livelihoods, poverty, or women and children. When thinking about this issue, it is worth considering if there are alternative sources of funding.