SCALING Up

Crosscutting Issues Affecting Scaling:

A Review and Appraisal of Scaling in International Development

Global Community of Practice on Scaling Development Outcomes

February 2021

Dr. Richard Kohl, Strategy and Scale LLC

Acknowledgements

This paper was written by Dr. Richard Kohl of Strategy and Scale LLC, on contract for the Global Community of Practice on Scaling Development Outcomes. It benefitted from multiple rounds of insightful comments and inputs from Johannes Linn. The author also received important and helpful feedback and suggestions from Larry Cooley, Laura Ghiron, Julie Howard, Dan McClure, Jenny Perlman Robinson, Sabrina Storm, Simon Winter, and Lennart Woltering. The paper represents the opinions of the author and does not reflect the views of the Scaling Up Community of Practice, its funders, or the Community's co-facilitators or members.

Table of Contents

A. Introduction and Purpose of this Note

The Global Community of Practice on Scaling Development Outcomes has now existed for six years. It has nine working groups (WGs), the majority of which are in sectors like Health, Education and Nutrition. One WG is devoted to a single aspect of the scaling process – Monitoring and Evaluation.[1] The Scaling Up Community of Practice (CoP) recently completed its 2020 (virtual) Annual Workshop. The workshop was comprised of a one- to two-hour meeting of each WG and three plenary sessions. The plenary session topics were: Scaling During the COVID Crisis, Adaptive Management, and The Role of Political Consensus.

The purpose of this paper is to identify the issues facing scaling in international development, which cut across multiple WGs, for which further discussion, exploration, guidance, and tools would be of interest to community members. While the primary audience for this paper is CoP members, it is our hope that this paper will also be of interest and use to other scaling practitioners in international development, as well as the whole constellation of organizations engaged in scaling: funders, implementers, innovators, national and local governments, and private sector actors. For each issue identified, this paper describes why addressing it could make a significant contribution to the field of scaling. It goes on to suggest how further exploration of each issue could be practically addressed. In most cases, we recommend organizing workshops or webinars on the issue; a paper that would take stock of the issue – including relevant existing tools, examples or case studies, tools, and guidance.

This paper is based on reviews of the Annual Workshop session videos and summary notes included in the November 2020 CoP newsletter. These can be found on the CoP website: www.scalingcommunityofpractice.com. Where helpful, the information from the Annual Workshop has been supplemented by other sources to flesh out either the issue itself, or the ways it could be addressed.

Overview

In reviewing the proceedings of the CoP Annual Workshop, it is striking how much progress has been made by the CoP and in the discipline of scaling over the last six years. New concepts have been explored or developed. Old ones have been fleshed out and clarified, and guidance, tools, and examples created on how to apply them. In proposing crosscutting issues, we want to explicitly acknowledge how much has already been done. Indeed, some of what we are proposing is a review, cataloging, and stocktaking of the progress that has been made. Other suggestions explicitly build on what has been done and how to take it to the next level.

Ten major crosscutting topics emerged from the CoP Annual Meetings. In some cases, these were explicitly stated on multiple occasions. In others, we have attempted to bring underlying questions or issues to the fore. In the latter regard, it appears that a “New Consensus” has emerged on how to think about and practice scaling. This includes a list of concepts with adjectives like adaptive, agile, iterative, and participatory; and nouns like systems, learning, and partnership. Indeed, many of these concepts are themes that run through all of the topics elaborated in the paper, which is why we have chosen to include several of them as topics on their own: Learning (with Monitoring and Evaluation (M&E)), Partnerships,

[1] See www.scalingcommunityofpractice.com for more information about the Global Community of Practice on Scaling Development Outcomes.

and Systems. One could argue about how new all of this is.[2] Regardless, it sounds new and different from the state of the scaling field when the CoP was first created, let alone when scaling began to receive growing attention 15 to 20 years ago. *Articulating and Elaborating the New Consensus* is the first issue discussed below as **Topic I**.

The second major topic is the issue of Systems, Systems Change, and Complexity, in terms of how these relate to scaling. Systems Change is a fundamental component of the New Consensus, yet that "consensus" conceals differences in opinion or perspective and lacks conceptual clarity. For example, some scaling practitioners describe scaling as a subset of systems change; others see systems change or systems "thinking" as part of scaling an innovation;[3] and still others see systems change as a stand-alone form of scaling. Complicating matters further, there is also a discussion of the differences between systems strengthening, change, and transformation. As this appears to be one of the central conversations within the CoP, at a minimum a debate and discussion of the topic, definitions and competing points of view would clearly be of benefit; greater clarity is needed.

Systems and complexity are not identical, but closely related. Complexity shows up in scaling in at least three ways: the problem being addressed, the system(s) in which that problem is embedded, and often the solution itself. Not surprisingly then, any of the approaches and tools contained in the New Consensus are similar or even identical to methods to address problems that are characterized by dynamic complexity,[4] or what are called "wicked problems" and "social messes." Given that there is consensus within the CoP that systems are central to scaling, regardless of differences in definition and opinion, then considerations of complexity are similarly important to a better understanding of scaling and how to do it. An exploration of the relationship between systems, complexity, and scaling would clearly be of benefit to CoP members. The issues of *Systems, Systems Change, and Complexity* are **Topic II**.

For many donor institutions, the New Consensus is becoming the dominant paradigm for understanding and talking about scaling. Given that for many CoP members the views of funders, the projects they fund, and the investments they make, play a major role in constraining or facilitating their work and that of their organizations this needs to be acknowledged as a huge accomplishment. Unfortunately, it is a fair generalization that this paradigm shift in thinking has not yet been reflected in the actual practice of most donors, especially that of official donors. **Topic III,** *Mainstreaming the New Consensus*, calls for exploration on what is behind the gap between donors' theory and practice, and what could be done to close it. Mainstreaming the New Consensus would be greatly facilitated by examples of procedures and processes that funders could adopt, apply, or modify to their own ends. This is also discussed under **Topic III**.

---

[2] To some extent it may have already existed in the early approaches to scaling and may be more about a shift in language and emphasis. Examples would be the ExpandNet framework and accompanying tools and guidance, such as Practical Guidance for scaling health innovations, Geneva, Switzerland: WHO, (2009) and Nine Steps for developing a scaling strategy Geneva, Switzerland: WHO, (2010) both downloadable at https://expandnet.net/tools/, the Management Systems International framework and tools developed by Larry Cooley with Richard Kohl and Rajani Ved. Scaling Up – From Vision to Large-Scale Change. A Management Framework for Practitioners. 3rd Edition (2016). https://msiworldwide.com/additional-resources/msi-scaling-framework, and the work and writings of Antraud Hartmann and Johannes Linn who were then at the Brookings Institution. See their Scaling Up. A Path to Effective Development. Wolfensohn Center for Development, (2007) The Brookings Institution. https://www.brookings.edu/research/scaling-up-a-path-to-effective-development/ and Scaling Up. A Framework and Lessons for Development Effectiveness from Literature and Practice. Wolfensohn Center for Development, (2008) The Brookings Institution. Working Report 5. https://www.brookings.edu/research/scaling-up-a-framework-and-lessons-for-development-effectiveness-from-literature-and-practice/. All of them identified the holistic, iterative, and dynamic nature of the scaling process.

[3] In this report, for reasons of brevity, we will use the term "innovation" to refer to the whole lexicon of "things" that get scaled – innovations, interventions, programs, projects, etc. It does not refer to scaling as systems change in its various versions.

[4] Relationships between cause and effect are not clear and are "distant" in time and space.

At the same time, donors and their activities and agenda comprise only a small part of the picture in terms of who is engaged in scaling in Low- and Middle-Income Countries (LMICs). Scaling is increasingly initiated, undertaken, or driven by domestic actors – especially national and local governments, or coalitions of local communities, NGOs and their partners. This highlights the need for these local scaling activities to inform and be informed by the principles, guidance, and tools of the New Consensus, i.e. to make learning from international experience available to local actors and the converse.

Topic IV is *Information, Monitoring and Evaluation, and Learning for the New Consensus*. The New Consensus has several implications for M&E that the M&E WG has recognized. The first implication is that it requires a whole new suite of M&E tools. New tools are needed to support, among other things: (i) feedback, learning, and adaptive management; (ii) adaptation and alignment of innovations and systems to each other; (iii) tracking progress in achieving systems change; and (iv) costing the multiple aspects of scaling and systems change (discussed separately under Topic VI). Second, a shift to the New Consensus implies that existing donor M&E frameworks for accountability will be less appropriate, and implementers will require greater discretion in running scaling efforts. Mainstreaming the New Consensus requires the development of new M&E systems that provide donors with the accountability they require. Finally, in regard to the growing number of local efforts at scaling, it would be instructive to look at to what extent existing M&E and Learning tools are relevant, or could be made relevant to these efforts, or whether different tools are needed.

Important progress has already been made in developing guidance and tools in most of these areas, e.g. real time monitoring and feedback of scaling efforts creating indicators of systems change. It would be helpful to catalogue and stock take these tools, with accompanying examples of their application. At the same time, more needs to be done; new tools, guidelines, and templates are urgently needed, particularly indicators of systems change and the extent to which M&E for the New Consensus can be combined or blended with M&E of the New Consensus for accountability purposes.

Topic V, *Risk, Uncertainty and De-risking*, is perhaps the major challenge to successful scaling, since the scaling process is characterized by a lot of uncertainty. This particularly arises after development of an innovation, an initial pilot, and proof of concept. A variety of questions arise about the viability of large-scale implementation with impact. For example, is there demand by customers or beneficiaries (and by the implementing institutions) for the innovation? How much will the initial rollout and ongoing implementation at scale actually cost? In cases of scaling through private sector pathways, will all the market actors in relevant value chains be able to make money?

Uncertainty about these and other issues is a major driver of the so-called “Valley of Death,” i.e., when scaling gets stuck between small scale pilot and proof of concept and implementation at scale. Actors that would be involved at scale are reluctant to commit until these uncertainties are resolved. In many cases there is a Catch-22 behind the Valley of Death; large scale implementation can’t occur until there is proof that an innovation or systems change is viable at large scale, and that evidence can only be generated by implementation at large scale. This section identifies these and other key challenges that involve how risk and uncertainty affects the scaling process. It suggests that de-risking may be akin to providing quasi-public goods, and explores approaches to de-risking, such as subsidizing or incentivizing early adoption until a critical level of demand is reached.

In many cases scaling cannot advance, let alone become sustainable, without a better understanding of **Topic VI,** *Costs and Economies of Scale and Scope* (ESS). It will be essential for both mainstreaming and applying the New Consensus for donors and practitioners respectively, to have a handle on the costs of its components, like trial and error, participation, etc. A major cutting-edge issue is identifying and

measuring the costs of the initial and ongoing costs of implementation at scale; especially adaptation of both the innovation and the systems itself, i.e. systems change.

Many costing questions relate to whether ESS are achievable and if so, how to achieve them. ESS are not easy to achieve for innovations that have multiple components, and are not products or services, e.g., processes, procedures, business models. This can be particularly challenging when implementation at large scale involves many and various implementing partners, several target locations, and different demographic groups, i.e., multiple different contexts. Existing costing tools are not designed to be easily updated at various levels of scale, or easily modified to take into account differences in context and availability of local resources.

The CoP and its members have made important advances in broadening, deepening, and disaggregating the various types of costs associated with scaling, but more needs to be done. **Topic VI** suggests addressing the following issues: (i) identifying the costs associated with good practices in scaling, i.e. the New Consensus; (ii) understanding how costs evolve over the process of scaling at different levels of scale and across different contexts; (iii) providing guidance on how to assess whether ESS exist for broad categories of innovations, if possible broken down by sector; and (iv) offer suggestions on how to uncover and/or achieve ESS in particular cases.

Topic VII, *Variance and the Robustness of Impact*, relates to several other topics (complexity and systems, risk, uncertainty, and de-risking), but is treated separately for expositional clarity. In most scaling efforts, the ability of an innovation to demonstrate proof of concept is largely based on its impact as measured by its average or mean effect. Yet for many adopters, be they producers and providers, or beneficiaries and consumers, the variance in impact is equally, if not more important because it is associated with risk. This is particularly true for variance over time. The risk-averse, “safety-first,” smallholder farmer is the classic example; but risk aversion is pervasive in the sustainable adoption of innovations in health, nutrition, and other sectors.

Risk aversion makes high-variance innovations harder to scale sustainably, a fact which is usually only discovered after the fact, as adopters stop using these innovations over time. Moreover, there is substantial anecdotal evidence that the impact of innovations that have high, multi-dimensional variance is not robust and is, in fact inimical to resilience, disproportionately affects marginalized communities or those with the lowest resources. This raises the questions of whether such innovations should even be scaled. It implies that producing data on the variance of impact should be a mainstay of proof of concept, from both ethical and practical perspectives, and highlights the fact that many impact evaluations do not generate that kind of information. This section suggests looking more closely at the role of variance and risks in outcomes, and drawing out the implications of those findings for impact evaluations and proof of concept, as a criterion in assessing scalability, and when deciding whether or not to scale a given innovation.

Historically, scaling has been largely a supply-side process driven by innovators, donors, or both. The New Consensus implies a shift in the approach to scaling that emphasizes partnerships and participation from early in the innovation and scaling process. **Topic VIII** takes this a step further. *Demand-Driven Scaling* (DDS) suggests that the logical next step is to have innovation and scaling be driven by the demand side. In many cases, this will be by those (large) domestic public and private actors in developing countries, that are expected to at least fund, if not implement, innovations at large scale. In other cases, this would be local actors working at the grass root level. In DDS, domestic actors, whether large or small, national, or local will not only be involved or even co-equal partners; *they* will decide what issues and problems are important, which innovations are needed, and how they are to be scaled.

The DDS approach to innovation will likely become of growing importance. This is both because developing countries will demand it and have increasing capacity and political clout to insist upon it, and because it will be more successful. Moreover, as noted under the New Consensus above, increasingly scaling is taking place at the local, or grassroots level, with or without the involvement or support of international partners. If, as many expect, that locus of scaling continues and increases in magnitude, there will be greater demand by those actors for international support and recognition, reinforcing national insistence for not just local ownership or buy-in, but to be in the driver's seat.

Further work in this area would start with cataloguing existing examples and identifying trends supporting this evolution, and it is probably happening more than is recognized. It would build on this stocktaking to provide good practices, along with guidance and tools on how to apply the DDS approach in practice.

Topic IX focuses on the issue of Digital Technology, often synonymous with ICT, or the more vernacular terms, Digital or Tech. The role of Digital in supporting international development has received enormous attention and funding, and this is reflected in the CoP Annual Meetings, particularly in the Agriculture and Health WGs, and the plenary session on COVID. One of the major applications of Digital to scaling is its ability to deliver and receive information to large audiences regardless of geographical distance, at near zero-unit costs as long as the platforms and infrastructure are in place. To cite but two examples, this has huge applications to the diffusion of innovation side of scaling, as well the last mile problem, e.g. delivering primary health or agricultural extension services. Digital has amazing potential to supplement, complement, or even replace traditional human delivery systems, monitoring, and improving the performance of human systems.

In reviewing the CoP activities and some of the literature, one thing that becomes apparent is how difficult it is to disentangle the different ways that Digital is affecting scaling; digital innovations *per se* versus digital mechanisms or platforms for scaling, versus innovations in digital scaling mechanisms and delivery platforms. The discussion also highlights the many challenges facing Digital, as the buzz has confronted reality, particularly around the issue of equity. As is the case in scaling writ large, to the extent that scaling either relies on existing infrastructure and delivery systems; or the skills of adopters and the coverage of those systems; or if the presence of skills is unequal; scaling usually reinforces or even exacerbates existing inequality. Nonetheless, Digital, particularly as a scaling mechanism or large-scale delivery platform, and as a means for systems change, clearly deserves additional attention from the CoP.

Topic X addresses the subject of Partnerships and Collaboration (PNC). It is almost impossible to overstate the centrality of PNC to scaling because it is implicated in so many aspects of scaling, especially as it is closely linked to participation and inclusion. To name but a few, PNC is a highly effective means of mobilizing resources for both going to scale and implementing at scale. By leveraging existing and diverse resources, it is arguably the most common and effective solution to the core challenge in most cases of scaling: reaching more and maintaining impact without a proportionate increase in resources. PNC is equally relevant to the challenges of complexity; solving and scaling the problem requires input from diverse perspectives. Finally, scaling, especially in its systems change dimension, inevitably involves, affects, or implicates multiple, if not all actors in a system, requiring their support, acceptance, or at least acquiescence. This usually cannot (and should not) be achieved without their participation and inclusion. Indeed, perhaps one of the areas that deserves further exploration by the CoP is to explore the roles of, differences in and interrelationships between inclusion, participation, cooperation, collaboration and partnerships in scaling. This would be not primarily from a conceptual or definitional perspective, though that would be a good place to start, but in terms of practice and action.

The rest of this note deals with each of these topics in more detail.

Ten Crosscutting Topics

I. Articulating and Elaborating the "New Consensus"

Numerous interventions during the Annual Meetings make clear that the thinking and practice of scaling has shifted from traditional development approaches over the last 5 to 10 years, to scaling frameworks developed in the 2005-2010 period. For many years, development approaches and scaling used what might be called a "social engineering" approach. This had clearly specified activities, intermediate outputs, outcomes, and goals articulated in a Results Framework (and accompanying project and annual workplans).

A number of new components of good practice in scaling were repeatedly mentioned throughout the CoP Annual Meetings. They suggest a significant and widening departure from the social engineering approach. Arguably many of these components were part of, or implicit in some of the early scaling frameworks and guidance. However, many now use language differently, receive more emphasis, have been made explicit, or have greater saliency. Taken together they represent a new paradigm or approach to scaling, a New Consensus. Specifically, a partial list of the components of a New Consensus includes:

1. Adaptivity, flexibility, and agility in the process of scaling, using feedback, learning, trial-and-error and iterative approaches; a focus on results rather than activities; and a script to be followed;
2. An even greater focus on political economy considerations and issues of incentives;
3. A holistic, systems perspective, shifting where appropriate from scaling innovations to focusing on systems change and scaling outcomes;
4. Emphasis on partnerships and collaboration, participation, and inclusion;
5. Build in scale and sustainability from the very beginning of the process, especially in the design, testing and proof of concept phases, including sustainable business models (financing); and either align with relevant large-scale systems or recognize the implications for systems change;
6. Designing projects with the understanding that scaling is a 10 to 15 year process, so that where projects have shorter timeframes, one project creates the foundation for the next phase;
7. Pay more attention to the challenges of adopting, integrating, and mainstreaming innovations into large scale systems in terms of the financial costs, organizational and incentives challenges, and the change management process needed;
8. Emphasize sustainability in multiple dimensions: impact, political, financial/fiscal, institutional, and environmental; and proactively anticipating and mitigating unintended or neglected long-term consequences;
9. Pay attention to the implications of scale for power dynamics, gender equity, and other equity considerations, such as coverage, impact, and the (often) higher costs of reaching marginalized populations (how to reach the "last mile").

To the best of my knowledge, the "New Consensus" has not been written up. I recommend this as a first step, accompanied by "how-to" guidance, examples, and tools. Such a document can outline the various key concepts and elements of the New Consensus, and perhaps provide a "compare and contrast" to existing practice. The paper(s) could also identify specific areas or topics where there are tensions, disagreements, or a diversity of opinion, or where a consensus has yet to emerge. Such a paper will be of great use to practitioners. It would *de facto* establish a standard or benchmark for good practice that will support advocating with donors and diffusion within implementing organizations. While the CoP should

not be in the role of explicitly endorsing specific approaches, merely publishing such a summary will carry weight.

Preferably, such research will be at a minimum, participatory; drawing on interviews of multiple scaling practitioners. Ideally, it will either be done by a committee, or team, or iterate between design workshops and one or two authors. Drafting a paper this way would likely generate a greater sense of ownership and buy-in, cohesion, and team-building among participants, (e.g., the leadership of the CoP WGs), and be more likely to generate insights.

II. Systems, Systems Change and Complexity

A recurring theme in the CoP WGs is that of systems and systems change.[5] The increased attention to these topics reflects a growing recognition that the scaling of many innovations is not "plug and play", and, more generally, that scaling innovations is only one of way to effect large-scale change. Successful scaling of innovations usually requires significant changes in the ecosystem, implementing organizations, policy enabling environment, value chain, etc., if impact and sustainability at scale is to be achieved.[6] Of the many issues prominent on the development agenda, e.g., the SDGs like Food Security, Hunger and Nutrition; Poverty, Inequality and Climate Action; both the problems and solutions are characterized by significant complexity. While systems issues and complexity are distinct issues, they are connected, complex problems, and their solutions usually require a systems approach.

Systems can be defined as the set of enabling conditions that exist at target scale (and implicitly at small scale or the pilot stage). Scaling, from the taking-an-innovation-to-scale perspective, requires making explicit which enabling conditions were necessary or desirable to achieving impact at the proof of concept phase; to what extent those conditions differ at whatever is defined as large scale; and ways to address those disparities where they exist. Resolution can occur either through modifying or adapting the innovation, changing the enabling conditions at large scale, i.e., systems strengthening, or change, or both.

Alternatively, if scaling is conceived as scaling outcomes, then systems change, or even systems transformation, is itself a form of scaling, e.g., policy change or reform. This highlights the need for care with language and greater conceptual clarity. Systems strengthening, change, and transformation are frequently used terms in scaling. Are these simply different formulations of the same underlying concept, different degrees along a continuum of change, or are they substantially different concepts?

A third way that systems show up in scaling is that they affect the choice of going-to-scale pathways and the viable options for large-scale implementation. As noted in the Education WG session, how scaling is managed in a centralized system might differ from how it is managed in a decentralized system. [7]

No doubt systems are relevant to scaling in other ways; it is clearly beyond the scope of this paper to list them all, let alone clarify each one and their relationship to the others. Hopefully this brief overview offers enough evidence to make the case for clarifying the various ways that systems and scaling do and do not relate to each other could be valuable to the CoP and its WGs.

---

[5] Foundation for this position is provided by Larry Cooley and Jonathan Papoulidis in a paper calling for an outcomes-based scaling approach that is adaptive, resilient, and politically-savvy, and which prioritizes popular support, pooled resources, and donor coordination. See e.g. Larry Cooley and Jonathan Papoulidis "Tipping the Scales: Shifting from Projects to Scalable Solutions in Fragile States," Development Journal, (2018) Washington D.C. https://link.springer.com/article/10.1057/s41301-018-0155-8.

[6] Correlatively, it may also be of interest to identify innovations that require stable systems to retain effectiveness and sustainability, i.e. little or no systems change, or for which outcomes are sensitive to systems change.

[7] See the Report on the 5th Annual Workshop of the Scaling Up Community of Practice. P. 10/CoP-Newsletter-19.pdf

Complexity is different, though related, to systems. As noted above, issues like Fragile States, Nutrition and Climate Change are inherently complex, and being explicit about how complexity affects scaling would be valuable. Moreover, as noted in the Introduction, even traditional sectors like Agriculture, Education, Employment and Health are increasingly seen as inherently multi-sectoral and complex, and require holistic systems approaches. Discussion of Agriculture and Food Security policy and programs almost always includes the issues of resilience, environment and climate change, and gender, nutrition, and the availability and affordability of nutritious food for the population as a whole.

Many of the process innovations and components of the New Consensus are identical, or similar to, tools and guidance found in the Complexity literature. For example, solutions to problems characterized by complexity, or those categorized as "Wicked Problems" and "Social Messes,"[8] depend on how the problem is framed; stakeholders have radically different world views and different frameworks for understanding the problem. The New Consensus emphasizes the collaborative, participatory nature of scaling, not only because it is ethical or moral to be inclusive – though clearly it is – but because the more diverse stakeholders are involved in the solution, the greater the chances of success, especially in terms of avoiding opposition from vested interests.

This suggests three areas for further exploration. The first would be to provide greater clarity on systems and systems strengthening, change, and transformation. It might start with illuminating the relationships and differences between the plethora of terms currently in use: such as scaling an innovation; scaling outcomes; system strengthening or change as part of the scaling process; systems change as scaling; and scaling as a form of systems change. It might also provide insight into whether, and to what extent, these different concepts and related approaches are relevant or necessary.

The second area would involve an exploration of the implications of complexity, wicked problems and social messes, and perhaps beginning with a literature survey. This could generate major insights, benefits, and tools for scaling.[9] The third area would examine the relationship between systems and complexity. It would examine the relationship between systems and the complexity of: (i) what is being scaled; (ii) the scaling pathway; and (iii) of the problem and context.

III. Mainstreaming the New Consensus

Closely related to Topic I is the subject of Mainstreaming the New Consensus. On several occasions during the plenary session on Adaptive Management, Larry Cooley asked presenters and participants why donors, for the most part, have not implemented the New Consensus. To my listening, most responses were prescriptive, using words like *need* and *should,* but not offering an analysis or explanation.

[8] One definition of a wicked problem is "a social or cultural problem that is difficult or impossible to solve for as many as four reasons: incomplete or contradictory knowledge, the number of people and opinions involved, the large economic burden, and the interconnected nature of these problems with other problems." See https://www.wickedproblems.com/1_wicked_problems.php. Others define it has having some or all of the following characteristics (a partial list): "They cannot be studied through trial and error", "these problems lack an inherent logic that signals when they are solved", "They do not have a definitive formulation or the way a wicked problem is described determines its possible solution" and " All wicked problems are essentially unique".

[9] While a detailed summary of the vast literature on complexity is beyond the scope of this note, a few items may suffice to justify why further research on complexity and scaling might be worthwhile. One comprehensive resource developed for the ILO and FAO is *"Knowledge Sharing for Change; Designing and Facilitating Learning Processes with a Transformational Impact*" (IPK approach). A different approach with substantial overlap is the Problem Driven Iterative Adaptation (PDIA) that was presented in the Fragile States WG. Marc Steinlin and Catherine Widrig Jenkins (2010) Knowledge Sharing for Change. Designing and Facilitating Learning Processes with a Transformational Impact. Facilitation Handbook. Cape Town, South Africa: Indigenous Peoples Knowledge.

Scaling practitioners know and emphasize that large organizations, divisions, and the individuals within them, even those that are NGOs or in the public sector, have other motivations or incentives in addition to altruism or working for the greater good. Their incentives often make them highly risk averse, i.e., against doing something new or pushing the limits of existing culture, practices, and procedures.[10] Combined with the inertia and homeostasis that characterize large organizations, these incentives pose major challenges for scaling efforts if not addressed. Scaling practitioners know how to identify the concerns, interests, and incentives of key stakeholders and figure out how to align them with scaling objectives.

Getting donors to adopt, support, and implement the New Consensus, i.e. mainstream it, is *de facto* a form of scaling. It is presumably one of great interest to CoP members and WGs. While institutional inertia is certainly a factor, it seems incontrovertible that a major explanation behind their failure to adopt and implement the New Consensus is that to do so is not aligned with the incentives of most donors.

While donors and the scaling efforts they fund and support are highly influential, as noted in the Introduction, scaling is and does also occur outside of these efforts at national, local and grassroots level. Decades of building of community organizations like women's savings and lending groups, NGOs, and more recently social enterprises combined with media reports of actual scaling suggest that such activities are increasing in number and importance. There is no precise accounting of how much of this type of scaling is occurring, and indeed such an accounting or at least a survey would be worthwhile. If we replace "mainstreaming" with diffusion of the New Consensus, the CoP could build on a survey and share and disseminate its knowledge and practice to these less formal but equally important actors and efforts. Whether a survey or a survey and dissemination, it would be vital for this be a two-way process of information exchange. These local efforts no doubt have generated important learning that could improve the scaling work funded by donors or otherwise being done at the international level.

Returning to the idea of mainstreaming, the CoP could provide a great service to itself and the whole field of international development, by providing an explicit, detailed analysis of the political economy and incentive questions that have inhibited mainstreaming to date. This could be accompanied by an examination of other challenges and obstacles. There has been some work in this area already, particularly the work of Johannes Linn and various co-authors. For example, Linn led a Brookings team that analyzed the status of scaling in IFAD, including the ways internal incentives, culture, and processes and procedures presented challenges (and opportunities) to facilitating greater scaling.[11] More needs to be done. Building on the work at IFAD might include analysis of other organizations that have tried, or are trying, to take on new approaches to scaling; including smaller, more nimble foundations that have the flexibility to be innovative.

Ideally, an analysis of the barriers, obstacles, and opportunities to mainstreaming the New Consensus could take three additional steps. First, it could outline a strategy to mainstream the New Consensus grounded in the political economy, incentives, and barriers analysis. (For an initial and partial list of

[10] Larry Cooley shared with me the anecdote about his experience in South Sudan – all of the implementing partners agreed that taking a flexible, adaptive approach was the way to get the best outcomes, but when they asked the USAID Mission Director and Chief Contracting Officer for legal cover to do so as it might involve divergence from their written contracts and agreements, both refused to provide such cover.

[11] Arntraud Hartmann, Homi Kharas, Richard Kohl, Johannes Linn and Barbara Massler, "Scaling Up the Fight Against Rural Poverty: An Institutional Review of IFAD's approach," Washington, DC: The Brookings Institution, (2010) https://www.brookings.edu/wp-content/uploads/2016/06/10_ifad_linn_kharas.pdf, his chapter "Incentives and Accountability for Scaling Up" in Laurence, Akio Hosono, Homi Kharas, and Johannes Linn, eds. Getting to Scale. How to Bring Development Solutions to Millions of Poor People, Washington, DC: The Brookings Institution (2013) and Johannes Linn, "How to mainstream the scaling agenda into funding and implementing organizations?" PowerPoint presentation to *Innovations in Agriculture: Scaling Up to Reach Millions,* Purdue University, 26 September 2017, https://ag.purdue.edu/scaleup/Presentation%20Library/Linn-Evening%20Discussion%20Group%20-%20Institutionalization%20of%20Scaling.pdf

questions that a strategy might answer, see Box 1 below). Secondly, it could consider who would implement such a strategy, or its components, as it is not clear that this is the CoP's role.

Box 1. Strategic Questions for Mainstreaming the New Consensus in International Donors

- What types of evidence, advocacy, marketing and communications would be needed to create momentum for mainstreaming, i.e., what would donor decision-makers need to see?
- What types of pilot projects or other "experiments" would be necessary to generate that evidence?
- Who would be the champions and leadership; with what resources of legitimacy, status, etc.?
- Does the New Consensus need to be broken down into a sequence of more manageable chunks to be adopted gradually or in phases? What would that look like?
- What would be the internal processes, procedures, rules and regulations that might need to be modified or replaced? How do these and other factors create institutional inertia that would need to be addressed? What is the perceived "cost" of such systemic and organizational changes?
- What donors, or divisions, or individuals within the donor community are likely to be first- or early- adopters? Who are the lead influencers or nodes in the donor social network that others are most likely to follow?
- Are there sectors that are more propitious as starting points, either because they've already begun to adopt mainstreaming, have an inherent affinity for this kind of process, or have failed to achieve their goals? This seems to be the case for Fragile States, COVID, and other emergencies such as in Afghanistan, where standard development practices are suspended. How can that experience be leveraged?

Finally, this work could develop sample or generic new processes, guidelines and templates for prospective donors interested in mainstreaming the New Consensus in their organization. For example, in the case of larger donors, what would an RFP, Results Framework, Reporting Requirements, Contract or Grant agreements, or project workplans look like if they were part of a New Consensus scaling project? How concretely could a donor align their standard 3-5 year project time frame with the fact that most scaling efforts take 10 to 15 years? How would a donor include, in a project's incentives, measurement and accountability for sustainability at scale? To the extent they exist, examples could be drawn from donors like foundations that may have made progress in institutionalizing the New Consensus because of their smaller size and greater flexibility.

IV. Information, Monitoring and Evaluation, and Learning for the New Consensus

Discussions of the New Consensus suggest it requires new systems for generating information and evidence. Two issues seem to be the most salient. The first is the need for new and/or additional forms of monitoring and evaluation (M&E) to support adaptive management and learning from trial-and-error. The second is that current systems for accountability are incompatible with the greater discretion needed to apply the New Consensus, and new systems are needed.

The M&E WG has already started to address these issues. The WG discussions noted, *inter alia*, that the New Consensus requires greater investment in indicators for learning and feedback at a much more granular level. Further research and development are needed to answer questions such as: what data is

needed about context, both at the beginning and to track changes? How do we track progress in creating systems change and sustainability in all its dimensions? How do we measure whether a five-year project has created the necessary foundations for the next in a sequence of scaling projects, if we accept that scaling will often require multiple projects over a 10 to 15 year period?

There were several interventions in the Annual Workshop sessions, which indicated that if implementing partners are to apply the New Consensus, they will need to have more discretion than is currently allowed in their contracts, grant agreements, or memorandums of understanding. At the same time, if donors are to adopt and apply the New Consensus, they will need guidance on how to combine discretion (and delegation) with their own fiduciary responsibilities, i.e. accountability. Providing donors and practitioners greater clarity regarding this challenge, if not resolution, and even models and templates of how to do it, will support mainstreaming of the New Consensus.

Research is needed in all three areas: M&E for applying the New Consensus; especially for *measuring progress* towards scale, systems change, and sustainability before having achieved big numbers; M&E for accountability of scaling efforts using the New Consensus; and blending discretion and accountability in the context of the New Consensus. This research would begin with a stocktaking of what tools already exist; which appears to be in process in the M&E WG.

V. Risk, Uncertainty and De-risking

Risk and uncertainty are omnipresent throughout the process of scaling and development. (The way risk shows up at a more micro level is discussed below under Topic VII, Variance of Impact and Robustness.) A partial accounting of the way risk and uncertainty show up in scaling yields three types. First, information uncertainties arise repeatedly in the scaling process and often deter external actors like the private or public sector from being involved and investing in it. As noted above, these particularly occur after the pilot and proof of concept phase, or initial scaling efforts, and are a major driver behind the Valley of Death. Examples of the information needed include: the existence of large-scale consumer demand and willingness to pay; the (in)compatibility of innovations with existing systems; the relevance and impact of innovations in different contexts; and the adaptations needed – or (re)confirming – that proof-of-impact at large or larger scale. In some sense, there is a Catch-22 in scaling, which is that many actors are reluctant to invest in large-scale adoption and implementation until an innovation's viability has been proven at large scale.

In most cases these uncertainties need to be resolved before further progress can be made. Donors interested in advancing the scaling process have found that if they want to move the process forward, they need to absorb some or all of the initial cost of information generation until other actors feel comfortable assuming a greater role. Generally, donors will need to play this role; either because information generation and de-risking activities are quasi-public goods; domestic actors are too risk-averse or resource-constrained; or both. In addition to financing the activities that will generate the necessary information there are a large and growing number of financial instruments and innovative contracting mechanisms that can be used to de-risk the activities of other actors and encourage them to participate in the process. These include insurance, first loss contracts, subsidies, and equity positions. A number of these were discussed in the Climate Change WG session.

Second, risk shows up in terms of project and institutional risk in development projects. Not only is risk central to advancing scaling, using a scaling approach can help donors address the risk they face in their own efforts. Linn and Chandy have argued that for Fragile States:

"…any attempt to use aid more determinedly to tackle the root challenges of fragility forces donors to accept a higher likelihood that project objectives will not be met and to adopt approaches that expose the donor agency to a greater reputational risk... Scaling up helps donors to … ambitiously tackle development risks without allowing institutional and project risks [by] employing greater selectivity in determining which interventions to scale up and … [a]dopt[ing] a portfolio approach, combining interventions with higher and lower risk."[12]

This leads to the third aspect of this topic, scaling innovations that themselves address risk. [13] The Climate Change and Agriculture groups both emphasized the importance of having donors underwrite the costs of acquiring the information needed for a financial instrument to function, e.g., insurance. Innovations that address risk and uncertainty (or promote collective resilience) arguably face the challenge of proving that they work at large scale more than most innovations because they are intrinsically characterized by economies of scale and scope (see Topic VI below).

Research, interventions and tools, and guidance are needed on all of these topics. Research on the first topic would look at:

- The risks and uncertainties inherent in the scaling process;
- Risks and uncertainties that are general, versus those that are specific to different sectors, types of innovations or scaling pathways, e.g., public vs. private;
- How these uncertainties affect the willingness of customers, clients, and other "retail" adopters to engage in the scaling process. This includes actors engaged in funding, production, distribution, and delivery of an innovation at large scale;
- Tools and methods for addressing these risks and uncertainties, and de-risking them. This would include looking at their pros and cons, what problems in the scaling process they address, and when and where they are appropriate;
- How de-risking can be done without creating perverse incentives, time-inconsistency, or otherwise adversely affecting long-term incentives for public and private sector actors;
- The extent to which resolving given uncertainties has the characteristics of public goods, like the lack of excludability; and where there is a case for donors to fully cover or subsidize the costs of resolving uncertainty.

This could be accompanied by research on scaling risk-reducing innovations themselves. For example, what can be said about the challenges and good practices in scaling crop or weather insurance, or innovations to improve resiliency? While it could be a separate topic, it may make sense to combine the two topics, de-risking scaling and scaling de-risking, or at least explore any overlap.

[12] Lawrence Chandy and Johannes Linn, "Taking Development Activities to Scale in Fragile and Low Capacity Environments." Washington DC: The Brookings Institution. (https://www.brookings.edu/wp-content/uploads/2016/06/Scaling-Up-Fragile-States.pdf)

[13] While these three aspects of risk and scaling – addressing uncertainties in scaling, scaling as a way of reducing risk, and scaling risk-reducing innovations – may appear to be distinct, the example of insurance serves as a counterexample. Insurance is used as a way of inducing actors to participate in the scaling practice by de-risking it, decreasing the risk that a project with other components will fail, and a tool that can decrease the risks for ultimate beneficiaries. Sometimes it does all three. The Climate Change WG focused heavily on these issues and particularly the application of innovative financial products and approaches like insurance, equity tools to crowding in private investment, address capital market failures, and mitigate political risk. For those interested, it is worth watching the video. https://www.scalingcommunityofpractice.com/resources/climate-change-working-group-meeting-october-30-2020/

Finally, better approaches to scaling could help donors address project, institutional, and reputational risk. Even though the New Consensus is probably more costly in terms of the initial upfront costs, it is likely to be much more successful in achieving sustainable scaling with impact. In other words, it will help donors achieve their objectives and reduce the risk of failure. Research in this area could contribute to getting donors on board with using the New Consensus and applying a scaling approach more widely in their activities.

VI. Costing and Economies of Scale and Scope

Costs are an essential component of the scaling process. A growing number of scaling efforts already measure unit costs and cost effectiveness;[14] this topic builds on this foundation by adding three more issues regarding cost. As the scaling field has evolved, it has revealed the increasingly complex way costs show up in that process. Examples include:

- Measuring tradeoffs between fidelity, quality, impact, equity, and cost;
- Unbundling innovations and interventions by component along measures of cost effectiveness, to allow for simplification, modification, and fitting intervention costs into existing fiscal space, or willingness to pay;
- Informing the tradeoff between adapting innovations to systems, versus adapting systems to innovations. In other words, measuring the costs of modifying, adjusting, or changing an innovation, and the costs of systems change such as increasing capacity, or aligning cultures and incentive;
- Providing large-scale adopters, particularly the government, in public-sector scaling, with estimates of ongoing operating costs after initial implementation.

Of these topics, the one that perhaps got the most attention in the CoP Annual Workshop was the issue of the costs of actually rolling out an innovation at scale, both direct and indirect. By direct costs we mean adding capacity in human resources, equipment or infrastructure. Indirect costs include the costs of adapting the system to align with the needs of the innovation so as to maintain impact, as well as the direct implementation costs, including adding additional capacity. Indirect costs could include changes in procedures, culture, processes, incentives, oversight and supervision, and other organizational changes; as well as a change management process to address potential opposition and resistance. Research is needed on how to measure these costs accurately and efficiently, perhaps accompanied by estimates of typical costs in scaling for certain types of innovations in given sectors.

The second issue is the cost of scaling using the New Consensus itself. There may be no comparative data on individual and collective costs of the components of “traditional” scaling, versus those of the New Consensus.[15] Research is needed on the costs of applying the New Consensus. Information on these costs, and ideally costs in terms of outcomes and probabilities of success and sustainability, would help mainstream the New Consensus with donors. For practitioners implementing the New Consensus, they will face repeated challenges, both in terms of their own decision-making and accountability to funders (or boards) to find the sweet spot between costs and benefits. In other words, identifying how much

[14] Unfortunately, many scaling efforts still suffer from lack of even this most basic data on unit costs or cost effectiveness. This is less and less because the tools are missing and more about increasing awareness of these tools and ensuring that they are being adopted and used. Nonetheless, it is important to recognize that progress is being made.

[15] See paper by Cooley and Papoulidis that argues for including explicit strategies for enhancing resilience and social capital as key elements of scaling in fragile states – Larry Cooley and Jonathan Papoulidis “Tipping the Scales: Shifting from Projects to Scalable Solutions in Fragile States,” Development Journal, (2018) Washington D.C. https://link.springer.com/article/10.1057/s41301-018-0155-8.

participation or trial-and-error is enough, given the costs. Identifying the sweet spots will support application of the New Consensus, make it more cost-effective, and once again support mainstreaming with donors.

The third area where more research on costs is needed is on Economies of Scale and Scope (ESS). As noted in the introduction, once scaling moves beyond simple products or interventions with few components that are largely plug and play, standard business assumptions of declining costs in manufacturing/production and distribution are of limited use. As Hartman and Linn noted:

> *"…costs of scaling up are specific to both the type of intervention and its particular setting (Johns and Torres 2005). The review concluded that a representative cost curve for health care cannot be constructed, due the lack of data and the difficulties to transfer cost estimates across settings."* [16]

The scaling community has done some work on ESS already, such as encouraging innovators to measure cost effectiveness and impact on a disaggregated basis to allow for evidence-based unbundling, scaling down for scaling, or scaling by subtraction. Another way the scaling community has addressed this issue is by suggesting innovators focus on delivery systems, business models, or both, that can then be applied to multiple product or service innovations, e.g., bundling. A third aspect concerns the extent to which innovations will need to be tailored, modified, or adapted to different contexts, and the cost implications of doing so. Such considerations are central to whether ESS can easily be achieved; the more customization needed, the greater the costs.

Recognizing the work already done, there is an opportunity to take achieving ESS to the next level. It is not just a new way of asking about disaggregated cost effectiveness to ask what within a package of interventions are the key drivers of cost and impact. Similarly, the scaling community would benefit from guidance on how to design models that are inherently more generalizable across diverse contexts, and therefore have economies of scale, or fewer diseconomies of scale. A quick review of the WG sessions yields mention of several methods and tools to achieve these objectives that have not been cataloged and have, therefore, received too little attention in the scaling community.

In sum, the issue of costs and economies of scope and scale are multi-dimensional. They have more depth, breadth, and complexity to them, in part because the New Consensus is a more sophisticated approach to scaling. The current treatment of both issues in the scaling literature has not kept up with the need, nor does it reflect this complexity. There is an opportunity to take stock of and catalog what has been done, and with additional research, to take this to the next level.

VII. Variance of Impact and Robustness[17]

Traditionally, the proof of concept phase of innovation and scaling has focused on measuring impact, along with cost effectiveness and efficiency. This invariably is a measure of *average* or *mean* impact, and in its more sophisticated versions, with some look at statistical significance.

Increasingly, for many scaling efforts, simply measuring the average impact of an innovation is insufficient for scaling. This is because adoption decisions are at least equally dependent on the variance of impact,

[16] Hartmann and Linn, op cit. p. 22

[17] The robustness of impact is a specialized topic that arguably could be included under Risks and Uncertainty, Systems Change and Complexity or even Costs and Economies of Scale. Given that all three of those topics already contain at least two or more subtopics, we have chosen to treat it separately.

perceived by many as risk, and many key scaling actors are highly risk averse. For purposes of this topic we will focus on internal variance, i.e. variance or risk that is at least partially under the control of the innovator or the adopter.[18] Internal variance is of several types: intrinsic, partial adoption, infidelity, input quality, and generic user ability.[19]

Many innovations and interventions have intrinsic variance. Sticking with the agriculture sector, think of a new seed. Even if it is planted in identical soil, with identical weather, fertilizer, water, care, etc., farmers will still experience some variance in yields.

Partial adoption and infidelity, and the variance they produce, have to do with the fact that many innovations, especially technologies, are a package of component interventions who interact with each other to produce impact. Hybrid seeds require specific (and greater) amounts of fertilizer and other inputs than open-pollinated varieties. They can also require adoption of other "good" agricultural practices like land preparation, planting, weeding, harvesting, etc. In many cases these are a change to, or departure from existing practices.

Experience has shown that many adopters, at least initially, adopt an innovation package sequentially, a few components at a time (partial adoption). For example, they may adopt the new seed and fertilizer but not change plant practices. Similarly, they often don't implement those components and practices accurately (infidelity) or fully (partial adoption again); they may use less than the recommended amount of seed or fertilizer or apply the fertilizer incorrectly. They may also make other choices that compromise quality such as using inferior fertilizer (input quality). Many of these decisions are driven by risk-aversion or resource constraints of one form or another, and usually the two interacting with each other.[20] (For the sake of simplicity, for the rest of this section we will use the term partial adoption to include sequential adoption, infidelity, or issues of input quality).

Partial adoption is not intrinsically problematic, it depends on what the drop off is in impact and how that affects the adopter experience. Is the drop off in impact rapid and steep, linear, or very modest for even significant departures from complete adoption, what might be called robust? Given partial adoption, are there optimal partial adoption decisions that will still produce significant impact and a positive experience for the potential adopter? To be specific, does use of 80 per cent of the recommended amount fertilizer, all else equal, result in 80 percent of potential impact? 50 percent? Even less? What happens at 70, 60, 50 per cent of the recommended amount, and so on?

The answer to both these questions requires data as to how outcomes are affected by partial adoption given the interaction effects between different components. Unfortunately, in many cases this information either doesn't exist or, when it does exist, is not provided to potential adopters. Without that information, adopters can't make decisions for optimal results and instead do so based on what may be cheapest or require the least amount of change in practices. If in fact drop-offs in impact are rapid i.e. impact is not robust in the face of partial or sequential adoption, then adopters will have a poor experience and not continue to use the innovation.

[18] What might be called contextual or external risk is a separate topic, e.g. weather. It is considered separately, at least to some extent, under Topic V, Risk and De-risking.

[19] We ignore variance associated with prices for simplicity's sake. Prices, whether input or output, and other forms of market risk, can be important, especially in cases of scaling through private sector pathways.

[20] These sources of variance can be caused by a variety of reasons, including conscious choices by adopters because of factors like risk aversion, resource constraints, ignorance about the whole package or its components, constraints on other inputs like family labor available at peak times, or traditional gender roles.

These issues are not relevant to agriculture alone. In health, nutrition, and other sectors, it is at least equally common that innovations are a package of interventions and these sectors too experience partial adoption, dropped components, infidelity, or compromises on quality.

Intrinsic variance and partial adoption have critical implications for sustainable scaling. Either individually or in combination, these sources of variance can mean that for many adopters the impact they experience may not be enough to justify using an innovation a second time or to continue to use it over multiple periods. In other words, the extent of scaling is less than it could or should be. This is particularly the case where adopters don't have information on how to optimize partial adoption.

The lack of data on variance also suggests that some innovations that are currently being scaled shouldn't because they are not robust over time or in the face of partial adoption. Yet many innovators and their funders continue to support scaling of such innovations out of ignorance; they neither measure intrinsic variance or robustness in their proof of concept efforts nor do they design innovations to minimize variance or maximize robustness.[21] In a period where development practitioners are increasingly concerned with the issue of resiliency, scaling such innovations seems at best at cross purposes with development priorities, if not counterproductive.

I recommend that the CoP support further research on Variance and Robustness. This might include:

- Developing a typology of variance and its underlying causes;
- Measuring intrinsic variance and robustness in innovations as they are being developed with the goals of minimizing the former and maximizing the latter. This could be combined with assessing permutations and combinations of components to allow for simplication, cost-reduction and adaptation;
- Including thresholds for variance and robustness in decision criteria regarding whether to scale an innovation;
- Producing information that will allow potential adopters to make informed partial or sequential adoption decisions and disseminating that information as part of scaling efforts.

VIII. Demand-Driven Scaling

Much of the scaling literature and practices focus on what might be called "supply-side" scaling. By this we mean going-to-scale that is driven by either an innovator working to scale their innovation, or an effort to affect systems change by a party outside of the system being changed. It is often linked to donor support for scaling of innovations whose development they have financed.

Over the last few years, the focus on scaling, as reflected in the New Consensus, has shifted away from a relatively pure supply-side approach. As noted under Topic I, good scaling practices include crosscutting partnerships, collaboration, and participation, and building all of these in from the beginning of the innovation and scaling process. This might be called "balanced" scaling – as in an equal partnership between innovators and funders on the one hand, and domestic actors expected to fund and implement at large scale on the other. To cite but one example, Village Reach and its partner organization Spring

[21] They also can't share this data with other actors in the scaling process who need to make investment decisions, such as parts of the agricultural value chain. They may find themselves producing or distributing a product for which demand is short-lived.

Impact are both promoting balanced scaling as good practice and have developed a methodology embodying that approach.[22]

These are welcome developments and to be lauded. Even though balanced scaling is still in its early stages the CoP may wish to anticipate what is likely to follow what might be called "demand-driven" innovation and scaling (DDS). This means scaling *and* innovation that is (or could be) led or driven by governments or private actors in developing countries, with innovators and funders acting responsively in support of their needs. While in principle it could apply equally to public or private sector scaling pathways, the case is clearly strongest for public-sector scaling.

While DDS is not necessarily the right way to scale in all, or even in a majority of cases, it has at least three strong arguments in its favor:

- Development funding (as in multilateral, bilateral, and foundation funding) is still relatively minor in terms of the funding for ongoing implementation at scale, compared with public expenditure and/or private financing;
- Almost all donor scaling strategies presume that some combination of (usually domestic) public and private actors will be responsible for funding at scale if scaling is to be considered sustainable;
- Many of the challenges that bedevil scaling – such as exit or hand-off strategies, generating buy-in and local ownership, aligning incentives, creating political will, aligning an innovation's needs with systemic resources, capacity, and culture – become moot issues when the very actors that will be funding and implementing at large scale are driving the process from the very beginning.

Box 2. Emerging Examples of Demand-driven Scaling

There are some early efforts to emphasize balanced scaling or another alternative, scaling through large social enterprises or growing such enterprises.[23] One example of a move towards balanced scaling is Grand Challenges Canada (GCC), which is funded by the Government of Canada as well as other funders, to support innovators in testing and transitioning to scale innovations that improve global health, humanitarian assistance, and health and economic prosperity in Indigenous communities. Many of its global health grantees who are scaling through the public sector have encountered significant obstacles, and their scaling efforts have not met expectations. In response, GCC commissioned a recently completed analysis[24] of the demand side of scaling, to identify the challenges and determine what can be done to address them and catalyze public sector demand. The solutions from this analysis will be implemented in a pilot effort in East Africa in early 2021. While this effort only involves demand-driven scaling (and not innovation), GCC's leadership has already signaled its commitment to integrate the demand side into the innovation process. This could foreshadow a gradual move towards demand-driven scaling and innovation.

[22] One organization active with this approach is Village Reach https://www.villagereach.org/ and its partner organization Spring Impact. See "The Journey to Scale with Government", https://www.villagereach.org/wp-content/uploads/2020/10/The-Journey-to-Scale-with-Government-Interactive-Tool_Final-2.pdf

[23] For more about scaling through social enterprises see important contributions by IMAGO, (https://imagogg.org/) the World Bank (e.g. Tinsley, Elaine; Agapitova, Natalia. Reaching the Last Mile: Social Enterprise Business Models for Inclusive Development. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/29544

[24] That report "Enhancing Public Sector Demand for, and Scaling of, Health Innovation, Landscape Analysis and Potential Pilot Model," is available upon request from the author, Richard Kohl,

A second example is the Millions Learning Real Time Scaling Labs being run by the Center for Universal Education at the Brookings Institution. While these labs initially focused on helping innovators or innovative projects and programs develop scaling strategies, they are now shifting their emphasis to work with governments, not simply as an actor to be "sold" an innovation, but to help them find or even co-create innovations that are aligned with their policy goals and objectives.[25]

Partnership, collaboration, and early participation already play a prominent role in the New Consensus and implicitly support balanced scaling. The long-term trend towards greater local/country control of the development agenda, and of innovation and scaling in particular, makes it likely that there will be political pressure to move beyond balanced to demand-driven scaling in the future. Research on demand-driven examples, beginning with a survey of examples and practices, would allow the CoP to get ahead of the curve and be a thought leader in this area.

IX. Digital

Digital has enormous potential in scaling. Digital is used to refer to innovations like apps, hardware platforms, and delivery systems; such as a website, mobile, or Internet. It is also used to refer to computing and/or data collection and monitoring technologies, or combinations of these. There is a general consensus that it has huge potential and it is a topic that deserves further exploration by the CoP and the scaling community in general. At the same time there are some significant concerns and downsides. We will address both in this section.

From a scaling and systems change perspective there are four particularly exciting aspects about digital. Given that the discussion of digital was most involved in the Agriculture WG session, we can illustrate with examples from that sector. The first is its use as a two-way mechanism to collect, deliver, and share information, training, and support. In agriculture, in most low resource settings, extension systems remain weak despite efforts to strengthen them. Digital is now being used to both strengthen existing human systems and to supplement them with virtual approaches and other technologies. One example is to share advice and information with farmers about good agricultural practices, such as giving extension agents, agrodealers, and even farmers access to online databases, providing information through two-way or push SMS, or the use of tablets loaded with information in the field.[26] It is also being used to provide tailored or customized information through call centers. Another example is monitoring and feedback to improve performance. Drones, satellites, and remote sensing devices are being used to both collect information on the status of farmers' fields and provide real-time guidance on activities like the best times to water, weed, plant, harvest, etc. SMS and, to a lesser extent, internet access are now widely used to provide farmers and other value chain actors with access to, and information about, inputs, services, and markets, including about input and output prices.[27]

[25] For a description of the Real Time Learning Labs, see Jenny Perlman Robinson, Molly Curtiss and Patrick Hannahan,"Millions Learning Real-time Scaling Labs Emerging findings and key insights," Annual Reflection Brief, June. (2020) Center for Universal Education at the Brookings Institution. https://www.brookings.edu/wp-content/uploads/2020/06/Millions-Learning-Real-time-Scaling-Labs-FINAL.pdf

[26] Digital Green has emerged as a leader in the use of videos, made by local farmers, of farmers, and for farmers, to teach each other and share good agricultural practices. See https://www.digitalgreen.org/global-impact/

[27] Indeed, one study listed the following agricultural functions that can be partially or wholly addressed by ICT: Awareness creation; Information dissemination; Promotional; Advisory; Knowledge sharing; Technology transfer"; Training; Facilitate market access; Credit and banking access; Input linking; Mass advisory; Business planning; Monitoring and Evaluation; Linking and partnerships; Collect and respond to farmers" feedback. See Jack Barber, Ellen Mangnus and Verena Bitzer, *Harnessing ICT For Agricultural Extension*; KIT Working Reports 2016:4. https://www.kit.nl/wp-content/uploads/2019/10/KIT_WP2016-4_Harnessing-ICT-for-agricultural-extension.pdf

A second exciting aspect about digital is its ability to create social networks for diffusion of innovations and for sharing of information about good practices. Historically, many scaling practitioners working on what was then called diffusion of innovation, used social network analysis (first developed by sociologists in the 1950s). Social network analysis initially began by identifying how information actually flows within a population, and then was used proactively to identify efficient ways of spreading adoption without having to contact directly all potential adopters. The efficiency and cost-saving of social networks have of course been amplified in the last several years by digital platforms like Facebook and Twitter. Given that secondary adoption (or diffusion) is the Holy Grail of innovation/scaling, digital social networks have substantially increased the use of this approach to scaling.

Third, in both of the uses of digital discussed above, and many others, digital is characterized by low unit delivery and operating costs, huge economies of scale and scope, and network externalities. This is especially the case where digital is being used to supplement or replace weak human delivery systems, e.g. agricultural extension or telehealth, where primary care is weak or unavailable. Low cost and cost effectiveness are central to its appeal; studies of Digital Green have estimated that its approach is ten times more cost effective than traditional appeal.[28]

Digital also has huge potential and is already being used in scaling as systems change. Examples of important one-off changes are digitization of health records and payments. Perhaps even more important are cases where digital is able to improve administrative, technical, or financial efficiency at the systems level on an ongoing basis. Oversight and supervision have been a chronic weakness of health delivery systems, especially in marginal areas and remote regions, leading to chronic problems like absenteeism, and contributing to low service quality and access issues. Many public health systems are now using or rolling out digital methods to monitor the performance of frontline workers and provide feedback. These applications have yielded important improvements in health performance and outcomes and have also been applied to agriculture and other sectors.[29] Once again, digital's huge economies of scale and scope, near-zero marginal costs, and low operating costs are in evidence.

Taken together, there is an important lesson in terms of innovations and scale. Digital's greatest strength is as a delivery system and platform; digital itself as a (scaling) innovation. It can be used as a vehicle for many different innovations in any given sector, as well as across multiple sectors. This highlights that innovations in delivery systems, business models, and financing mechanisms that facilitate distribution and scaling have much greater impact than innovations that are to be delivered.[30] How to identify, build out and promote innovations in scaling mechanisms is an important area that the CoP could discuss.

At the same time, digital is not without its problems. First and foremost is that more emphasis has been placed on scaling digital innovations, e.g., smartphone apps and software, than on using digital for scaling. This has led to an explosion of hackathons and other digital innovations targeted to global poverty, but too often ungrounded in knowledge of the context, and other ethical and responsible practices in innovation.[31] A similar blind spot that neglects the role of the user, is that scaling of, by, or via

[28] Mark Bell. "ICT – Powering Behavior Change for a Brighter Agricultural Future" Information and Communication Technologies for Agricultural Extension and Advisory Services. Modernizing Extension and Advisory Services (MEAS) Discussion Report. October 2015.
[29] One such example is the CommCare mobile data collection platform, see https://www.dimagi.com/commcare/ The CommCare app has also been applied in agriculture, USAID's Nataal Mbay program uses CommCare to track inputs and productivity, manage loans, and collect rainfall data. CommAgri, Naatal Mbay's digital tool, is used by 55 cooperatives and 500 extension agents, who have registered 68,000 farmers."
[30] In that regard, another aspect of digital relevant for scaling is as a source of funding for startups, social enterprises, etc. Not only are funding sources more transparent and easier to find and access for innovation and scaling, but crowdfunding and grassroots fundraising have allowed innovators to generate financial support for scaling beyond reliance on traditional donors and foundations.
[31] For two sources of good practice and ethical principles in digital innovations, see https://digitalprinciples.org/ and the Humanitarian Innovation Fund's https://higuide.elrha.org/toolkits/get-started/principles-and-ethics/.

digital usually requires changes in attitudes, incentives, culture, and behavior by users, providers, or both. Returning to the example of applications to supervision and oversight of health workers, digital monitoring faced initial resistance by frontline workers because it was seen largely as a way of increasing accountability and thus punitive in nature.

> *A persistent challenge in resource-limited contexts is that the introduction of digital innovations may add increased scrutiny of performance and efficiency, which may also uncover institutional dysfunctions. Whether fueled by inertia, resistance to change, conflicting priorities, lack of training, absence of engagement, or lack of clarity in roles and responsibilities, it is important to consider such factors and understand them, to adequately plan for scale.*[32]

In the cases where such innovations have been successfully introduced and scaled, it has usually been with explicit emphasis on using it to improve performance rather than solely as a means to improve accountability accompanied by punitive measures. The focus on improved performance allowed initial resistance to be overcome because, by producing better health outcomes in their communities, health workers' motivation shifted from extrinsic carrots and sticks to (increased) intrinsic job satisfaction.

Second, in many cases digital has been seen as a panacea. To give but one example, while digital tools and platforms for education and training now play an important role in the learning field, it appears that they cannot and should not entirely replace hands-on learning. It seems we don't yet understand the limits of virtual, and where the sweet spot is between no-touch and high touch in terms of cost versus learning. As the experience with virtual learning under COVID has vividly demonstrated, no-touch cannot fully replace human interaction in the education sector itself, or in education and training in general.

Thirdly, despite the low operating costs, the ongoing costs of maintaining and updating websites, technical information and having actual humans staff call centers, and advisory services (even if delivered virtually) are not small and need to be accounted for. To the best of my knowledge, ICT and call-center delivered agricultural advisory support has never become financially sustainable on its own, and while efforts to bundle it with other services have led to significant cost-recovery, even bundled it has never reached breakeven levels, let alone profitability. In general, despite the attempts to leverage partnerships with mobile providers and others, scaling and delivery through digital means has not included a viable, sustainable, business model.

Fourth, while in certain ways digital innovations and scaling facilitate solutions to issues characterized by complexity, in others they don't. SMS, Twitter, and other technologies require reducing complex and multi-dimensional problems to byte-sized points. In many cases, this comes at a cost, including addressing only the needs of the average, or majority of users.

Finally, and perhaps most importantly, while the term digital divide is less widely used than a decade ago, the issue of equity remains as important as ever. Internet coverage has not kept pace with mobile coverage, and 5G coverage is still a long way off in many places.[33] Today, it is less about coverage and more about either the ownership and affordability of hardware platforms, e.g., smartphones; or who does or doesn't have the skills, education, and complementary resources in place to leverage digital

[32] Alan Labrique et al. "Best practices in scaling digital health in low- and middle-income countries." Globalization and Health. 14:103. November (2018). **https: //globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-018-0424-z**

[33] As of 2019, 5G was only available in Lesotho and South Africa in sub-Saharan Africa. As of 2020 it was not yet present in India, let alone in rural areas. Because 5G requires more frequent and dense installation of towers, there are even fewer incentives to undertake that investment in rural or remote areas with sparse populations or low purchasing power.

innovations effectively.[34] Marginalized populations are often older, less educated, and have other characteristics that render them less likely to either adopt digital innovations; to be able to use them effectively; or both. It should come as no surprise that in many efforts to use ICT as a delivery platform, radio and television have proven more effective and achieved greater scale despite their age and limitations as a one-way technology.

The same problem occurs for workers using digital tools, such as in digital health or health informatics. LMIC governments, with donor support, are now investing heavily in these areas expecting reduced inefficiencies, improved access, reduced costs, and improved health outcomes. Yet what is now being called the "digital health skills gap,"[35] or an "endemic lack of skills in information and communications technology," has emerged as principle obstacle to sustainable scaling. This is reinforced by weak ICT infrastructure within public systems, and a lack of sustainability; initial investments in network infrastructure and hardware are often donor-financed without the necessary domestic financial resources in place for support, regular hardware and software updates, and addressing normal depreciation. While this problem is peculiar to digital, the point is that digital is not an exception.

In sum, digital offers huge promise for scaling, and for good reasons: low unit costs, economies of scale and scope, substitutes for, or supplements to, human resource constraints and weak delivery systems. In many applications it is either an innovative mechanism for scaling, or a form of system change and strengthening. However, it is also beset by numerous challenges that are still not being adequately addressed. The CoP could make an enormous contribution to better understanding the role of digital in at least three areas: (i) untangling the current and future roles of digital conceptually and practically in terms of scaling; (ii) assessing its actual impact and potential, illustrated by examples; and (iii) exploring equity issues.

X. Partnership and Collaboration

Partnerships and Collaborations (PNC) are clearly central to scaling, regardless of sector, approach, or pathway. There are two reasons for this: first, scaling just about always involves systems change in some form; indeed, some argue that scaling *is* systems change. (See Topic II above). While there are no doubt systems composed of one organization or actor, most have multiple actors. So, systems change involves some combination of persuading, coopting, or coercing those actors to agree to the change. For political sustainability, coercion is the least desirable, so inherent to any system change is some form of collaboration. This is true not only of the outcome of systems change, but the process of getting there. Most actors who drive scaling don't have the resources – be they status, legitimacy, political power, access, or financial – to get agreement from all actors by themselves. They require allies and coalitions – synonyms for partnerships and collaboration.

The same is true for scaling up innovations. While there are extreme cases where one organization has all the resources, capacity, and capability to go to scale and implement at scale, this is rare. Even in cases of public sector scaling or its mirror image, commercialization; the necessary roles, responsibilities, and resources that have to be met are beyond the capacity of one actor on their own. This is especially the case when considering the quality or specialized resources, and in the aptly named low resource

[34] As discussed in the agricultural WG session, extensive evidence shows that digital innovations can easily hit a wall as the number of downloads of an app or software may generate only 20 percent ongoing engaged users; not a high success rate. This highlights the need for addressing the enabling environment to address equity and other considerations.

[35] See for example, the Measure Evaluation presentation on "Health Informatics Skill Up for Scale Up" presentation, https://www.slideshare.net/measureevaluation/health-informatics-skill-up-for-scale-up/measureevaluation/health-informatics-skill-up-for-scale-up

environments. In the plenary session on Fragile States, several speakers emphasized that the inherent scarcity of political and operational resources, quantitative and qualitative, necessitates a collaborative approach to scale as the default approach. PNC is arguably the best way to address the resource constraint, especially in terms of infrastructure or human resources.

Whether quantitative or qualitative or both,[36] this is one of the reasons that Public-Private partnerships (PPPs) have become one of the most common mechanisms for scaling; they are, by definition, partnerships. Indeed, in the Youth Employment WG, PPPs were mentioned so frequently that they appeared to be the default approach to scaling. In fact, it is probably the case that PPPs are the most common mechanism for scaling, even though they may not be called that, or formally or legally qualify as such. Five case studies conducted for USAID's Bureau of Food Security on successful scaling of agricultural innovations through commercial pathways, concluded that the public sector role was essential, even though none of them involved formal PPPs.

> *In most of the case studies, the public sector played a critical role in scaling, ensuring commercial sustainability, or both.* ***The importance of public sector involvement lies in ensuring buy-in to, at a minimum, avoid political interference or opposition****. Therefore,* ***the involvement of the public sector is necessary …****, even at the initial stages. [Emphasis in the original].*[37]

The importance of PNC in scaling and systems change has been recognized for many years, see the work on Implementing Policy Change from the 1990s and the MSI Framework (2005), both cited above. More recently, a study of scaling up in the United States, presumably a high-resource context, concluded:

> *None of the lead partner organizations we studied had all the resources and knowledge needed for scale up. They relied on partners to help them meet these needs. Across all pathways, supporting partners were identified as being very important.*[38]

Perhaps the most important is participation and inclusion; it may be theoretically possible to practice inclusion and participation without PNC, but in practice it is difficult to imagine. Moreover, to echo the discussion of systems, scale, and complexity above, solutions to complex problems can neither be identified nor scaled without having the diverse perspectives of all actors in the system involved. Inclusion is not only good practice from an ethical or political standpoint, it is intrinsically necessary to scaling to understand local context, the user's perspective, and complexity; all reasons for combining inclusion and PNC in scaling. In the Education WG session, inclusive partnership with parents is now recognized as essential for scaling, and that could probably be extended to other key stakeholders, for instance, educators.

[36] Indeed, one way of implementing at large scale is to work with a network or collaboration of actors who either all have the necessary capabilities and resources but lack large-scale capacity, or alternatively, a network of specialized institutions that combine their distinct capabilities but can all have the capacity to deliver at large scale. (And permutations and combinations thereof).
[37] See Richard Kohl, Colm Foy and Gwynn Zodrow. Synthesis Report. Review of Successful Scaling of Agricultural Technologies. Prepared by Management Systems International for the E3 Analytics and Evaluation Project. February, 2017 https://www.agrilinks.org/sites/default/files/resource/files/BFS%20Scaling%20Cases%20Synthesis%20Report%20508%203-9-17.pdf
[38] Diffusion Associates. *Strategies to Scale Up Social Programs*. The Wallace Foundation (2017). https://www.wallacefoundation.org/knowledge-center/Documents/Strategies-to-Scale-Up-Social-Programs.pdf

Box 3. Underutilized Approaches to Partnership and Scaling

There seem to be many approaches to PNC that are popular in the U.S. domestic context, but less widely used or known in international development. To give but one example, one collaborative approach for scaling is Qualitative Improvement Collaboratives (QICs). While they have been widely used in the U.S. and to a much lesser extent internationally, in both cases this has largely been confined to health. QICs combine innovation and scaling into one collaborative package, (*de facto* using the wisdom of the crowd or crowdsourcing), in this case the wisdom of collaborative participants, to design, test, assess, and scale innovation.[39]

The Wallace Foundation study, cited above, developed an explicit typology for partnerships[40] in the context of scaling U.S. domestic social programs, but those distinctions, and their implications for practitioners are neither well-known nor widely used in international development. There is no doubt that others are used by local actors in other OECD countries and domestically within LMICs that have also received little attention, let alone application in international development.

There are a large number of approaches to partnerships and collaboration that could be used to support scaling in international development but currently aren't. The discussion in the Agriculture WG suggested that there are private sector models of partnerships and distributed leadership that have yet to be brought to scaling. Exploring these possibilities could build on existing conversations within the CoP about the role of boundary-spanning intermediary organizations as conveners, organizers, and managers of collaboration; and the challenges they face; including finding sustainable sources of financing. Like many things in scaling and development, the organizational infrastructure for managing collaborations usually only lasts as long as that of the project providing funding.

Partnerships and collaboratives are more important than ever as the challenges that scaling is trying to address have become more complex and urgent. The likelihood of finding the necessary resources to either go to scale or implement at scale in one organization has decreased dramatically. For all of these reasons, PNC should be viewed as foundational and a core practice to scaling. There is an opportunity for the CoP to identify, assess, adopt and disseminate approaches to PNC that have yet to be applied to scaling in international development.

[39] In a QIC, teams from participating organizations collectively identify a problem, brainstorm on multiple solutions worth trying, and agree upon measures and metrics to assess progress on results. Each team implements the set of proposed solutions that make the most sense for their organization and context and share the results at regular intervals. After the collaborative ends, teams summarize their results and lessons learned, and those interventions that have proven effective are often widely adopted across participating organizations, as well as shared with non-participants. This approach embeds scaling into the innovation process. The leader in QIC is the Institute for Healthcare Improvement (IHI), see http://www.ihi.org/

[40] Diffusion Associates, op cit. This study identifies within what it calls the "overall partnership constellation," defined as all organizations working together to scale up a social program …[four types of partners] (1) Lead partners that direct scaling efforts, (2) distribution partners that provide connection to local implementers, (3) supporting partners that provide expertise and/or funding, and (4) implementing partners that provide direct services to intended beneficiaries," p.5. While the relationship is not quite a tight fit, this typology seems congruent with the Spaces and Drivers of the Hartmann and Linn approach, or the various roles and responsibilities involved in either going to scale or implementing at scale

Conclusions

This paper has identified and described ten crosscutting topics about scaling in international development that merit further exploration. They were selected drawing on the discussions that took place at the CoP Annual Meetings; in some cases by explicit mention, and in other cases as underlying issues. These issues are based on the author's subjective assessment of their importance, interest, or where further exploration might be most beneficial to CoP members and scaling practitioners writ large. No claim is being made that these topics are the ten most important to scaling at this moment, or that other topics are not important.

Beyond simply identifying each topic, this report has attempted to describe the parameters of each issue and suggest what a further exploration and investigation of each theme might look like. In the course of doing so, it became apparent how in several cases the issues themselves are so knotty and entangled that our recommendation for a first step is to get some definitional clarity, and clearly identify the different opinions on the topic. The prime example is the cluster of issues around Scaling, Systems Change and Complexity.

At a minimum, we hope that the topics identified resonate with enough CoP members that they, and hopefully other readers, are motivated to explore them further, and that the description is sufficient to serve as a basis for that exploration. Such exploration might begin with CoP workshops, webinars, or debates about these issues. More ambitious would be for the CoP or its members to write or commission papers on the issues discussed herein. Depending on the topic, these could range from concept papers to stocktaking that catalogs existing tools and guidelines, identifies gaps, and proposes or develops new or additional tools to address those gaps.

Four things are clear from this exercise. First, scaling is more important than ever to solving the issues of international development, and that is now recognized by a large and growing number of people and institutions. Second, scaling has come a long way in the last twenty years in terms of our understanding of what is good practice; and the development of tools and guidelines to translate that understanding into action. The depth, breadth, and sophistication of the topics identified in this paper stand on the foundations of those accomplishments and deserve acknowledgement. Third, there is a broad, rich, and deep ongoing conversation about scaling. In regard to both the progress and current conversation, the CoP has clearly played an important role, and should continue or even expand its activities. Finally, the topics identified in this paper clearly show that there is more work to be done, and while daunting, the task is also tremendously exciting.