Why we need to invest in the infrastructure for scaling, not just the product or organisation being scaled

by Warren Ang and Yanni Peng





By 2015, Huiling had become one of China's largest grassroots non-profit organisations spanning 14 provinces, 20 cities and serving 1,400 adults with intellectual and developmental disabilities (IDD) daily. They provided services such as group homes, daycare centres and supportive social services for education and employment.

However, the problem size was still massive: only 1.6% of the 7.3 million adults with IDD in China had access to the types of services it provided. Huiling was the largest service provider, and yet only covered 0.02% of the total market of 7.3m adults with IDD in China.

Facing this gap, in 2015, Huiling's leadership made a bold decision to focus on scaling their impact through replication. Instead of being a large service provider, Huiling would become a 'replicator/franchisor' of its core product consisting of its daycare centre and group homes, so that other community-based NGOs could adopt and scale this product to increase access to services for adults with IDD.

Through a series of tough work - redefining strategy, reorganising headquarters, revamping fundraising efforts, and standardizing a product package - Huiling managed to increase its growth rate of local franchisees by five times. Where it had only increased at a rate of one Huiling per year for 20 years, 4-5 new franchise 'Huilings' per year were created between 2016 and 2019, increasing its total to 38 cities across 23 provinces.

However, despite the significant improvement by taking a replication approach, Huiling as a single network of organisations was still only scratching the surface - 98% of adults with IDD still lacked access to services.

The problem: why we need to build the infrastructure for scaling, not just the product being scaled

This story is not at all uncommon. So what else is needed for us to solve social problems more quickly and sustainably at scale? Replication is well established as a strategy for scaling impact. The respective methodologies, insights, and challenges can be found in Bradach's 'Going to Scale' (2003), or Dees and Anderson's 'Scaling Social Impact' (2004). However, the dominant approach implied by the literature and practice thus far has been to drive replication 'one organisation or one model at a time.' We argue that this approach alone is not getting us far enough, fast enough, to solving social problems at scale. We need to invest in the infrastructure for replication, not just the product or organisation being replicated.

To understand why, one way to look at the challenge is that scaling impact requires success in both "upstream" and "downstream" activities. "Upstream" efforts look at scaling from the lens of the solution or intervention required to solve a specific social problem. It involves taking the interventions that have worked, and preparing them for replication (productization, preparation of standardized operating procedures, working out the governance, financing, capacity building and M&E models for replication etc.). "Downstream" efforts look at scaling from the lens of the population, community or delivery system the intervention will be integrated into. It involves developing operating models and systems that can meet the needs of local communities, and raise resources to sustain and scale delivery of the interventions.

Typically, funders and non-profits focus more on "upstream" than "downstream", yet "downstream" is where most scaling efforts fail. Good solutions are identified to specific problems, and funders and practitioners ready them for replication and scaling, but tend to underestimate the challenges of developing (often last-mile) delivery systems. Without strong "downstream" integration, products and interventions may never reach their target market at a meaningful scale, and local organisations that replicate the model may lack the capability or environment from which to raise resources, making

them unsustainable. Replication ends up slowing to the limit of the network or dies off when the central funding runs out.

Others, including us, have given thought to how downstream efforts can scale and be sustained by clarifying answers to the question of who pays and who implements at scale and, in "What's Your End Game?", we outlined several "end-games" for non-profits to consider. If it's more of a public good, we argue non-profits should target "government adoption" where the government would pay and either the government or civil society would implement. If the product has a market that's willing to pay, we would argue "commercial adoption" would be most efficient, where businesses implement the product and their revenues pay for it. If the product could be sustained by philanthropic contributions, we argue for "sustained service" or "civil society adoption". This is helpful framing as a target, however, the hard part is building the operating models and infrastructure that make these end-games possible in practice.

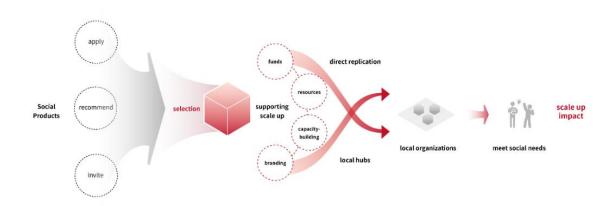
How to invest in the infrastructure to scale and sustain replication

Instead of only driving replication one organisation or model at a time, we can also invest in creating the infrastructure in a given social sector ecosystem to scale and sustain replication to the last mile. This infrastructure consists of building two "aggregators": 1) neutral hubs located at the provincial/state, city and/or county levels that can aggregate the local organisations and needs of the local community to drive adoption of products, source locally created products, and build the capacity of local organisations to mobilise resources from local government, corporate and public donors; and 2) a nation-wide product platform that aggregates the best products (models and interventions) and facilitates their replication.

The Effective Philanthropy Multiplier (EPM) in China is an early, but impressive example, of the exponential scale that can be achieved with the infrastructure to connect both 'upstream' and 'downstream' efforts in replication. EPM has created a network of 38 hub organisations at both provincial and city levels, connecting EPM to over 13,000 local

organisations. Its nationwide product platform has sourced 49 products or intervention models for different social service needs including disabilities, elderly care, child protection and environmental protection. EPM conducts roadshows together with the local hubs to facilitate the replication of these products to tens of thousands of local organisations. The graph below illustrates how EPM's product platform works:

Figure 1: How EPM's product platform works in collaboration with neutral local hubs



In just three years, EPM has managed to achieve 49,477 distinct replications through 8,607 local community partners, including local NGOs, community groups, schools, corporates and volunteer groups. Collectively, in these three years, the replication infrastructure, together with the hard work of the brand and local organisations, has reached 65m beneficiaries and mobilised an estimated 923m RMB (130m USD) of funding.

To date, based on EPM's monitoring efforts, the vast majority of replications have been sustained. This is in large part due to the ability of the hubs to help local organisations leverage the products to mobilise local government and corporate resources.

For example, in Guizhou, a children's service centre model developed by a leading foundation, was adopted by the local hub. The hub trained local organisations and replicated the centre in eight locations during the first year. The hub was familiar with Guizhou Government's priorities and needs, and was able to develop a partnership with Guizhou Government's charity foundation (a quasi-governmental foundation attached to the department of civil affairs), and a leading internet corporation to further adapt and scale the children's service centre to reach more remote and poor beneficiaries. The tripartite collaboration was established with joint funding of 4.84m RMB (680k USD) to build 17 more centres in the province and increase the annual operating funding for each centre from 120k RMB (17k USD) to 280k RMB (40k USD), which greatly improved sustainability and capacity. The following year, the provincial government committed another three years of funding to build 354 new centres, with a first year commitment of 16m RMB (2.3m USD) for 100 new centres. This is an example of the accelerated scaling made possible through partnerships with a local hub, which played an essential role in adapting the product to the local context, mobilising local resources, and training and capacity-building NPOs to implement it.

The exponential increase in direct impact that EPM has achieved shows that investing in the infrastructure for replication (neutral hubs and a product platform) can be a scaling impact game-changer. Hubs help to solve the downstream problem of scaling, primarily because they are neutral and can amplify many non-profits and can leverage local resources (local government, corporate and public), rather than remaining reliant on resources nationally. The product platform helps to solve the upstream problem: aggregating the best products/models for different social services in one place to make it easy for hubs to select models that work.

Large countries with a huge supply gap for social services at the last mile could adopt an EPM-like approach to replication. Our goal in this case study is to share our early experience with EPM to inspire philanthropists, foundations and non-profits to consider investing in the infrastructure of replication in their contexts. In this case study we will explore the origins of EPM in China, how it works and how it goes beyond traditional product platforms, the challenges and limitations of EPM and areas for improvement in

China, and thoughts on how EPM thinking can already be seen in other countries, using India as a starting example.

The origins of EPM in China

Narada Foundation is one of China's leading foundations focused on supporting the development of the country's social sector¹ and has led the movement for scaling impact, spearheading the development of EPM. When Narada Foundation did its strategy planning in 2015 and 2016, they identified three big obstacles to the potential of the Chinese social sector to scale social services to the last mile.

First, driven by government policy, China's local organisations grew from 506,785 in 2013 to 801,083 in 2017² (where nearly 690,000 organisations were city-level or county-level NPOs³). Many of these organisations wanted to address social problems but lacked the know-how to effectively deliver social services.

Second, after 10 years of rapid development, there had emerged a pipeline of 'brand organisations' (organisations which had created effective products or models that had expanded their coverage to dozens of cities or provinces). However, these brand organisations struggled to scale and replicate their product or model faster and further without either the channels and infrastructure to tap into new locations, or the support to find the local organisations and local funding needed to scale up.

Third, the Chinese government increasingly sought to outsource and pay for social services to local non-profit organisations, corporates were increasingly setting up CSR activities to fund social services, but both the government and corporates lacked effective and efficient means to identify the best models to support.

¹ Since 2007, Narada has been instrumental in setting up the social sector's infrastructure, including China Foundation Center, China Social Enterprise and Impact Investment Forum, China Donors Roundtable, and the Gingko Fellow Program dedicated to long-term investment in young NGO leaders.

² "The Bluebook of Philanthropy: The Development of Philanthropy in China (2018)", Yang Tuan, Chinese Academy of Social Science

³ "The Number of Social Organisations in China Surged to 800K+, Most Are County-Level", Wang Yong, Philanthropy Daily, December 2018

EPM was initiated to build the infrastructure and partnerships to address these interrelated gaps. EPM would support upstream development through the product platform and downstream development through building a network of neutral hubs. Additionally, EPM would facilitate collaboration upstream and downstream by mobilising new resources and partnerships to build the overall ecosystem. The EPM initiative gained support quickly - 16 organisations became core partners, including other leading foundations and top tech companies in China.

How EPM Works

EPM's team manages its work at these three levels. For upstream product development, brand organisations with products/models accepted by EPM are given capacity building support through EPM's 'Scale-up Academy' (a capacity building program), a scale-up grant to support the replication process, and networking connections to EPM's larger network of funders and service providers.

For downstream development, EPM provides a small stipend to each of the 38 local hubs for their work and contribution to the process of replication. Together with each local hub, EPM arranges local roadshows, where based on local needs, certain products/models from the platform are presented, and EPM and the local hub helps to match products/projects with local organisations. Once matched, the brand organisations and the local organisations work out how to move forward and develop their own financial relationship based on the local situation. Local hubs, if capable of supporting the process, continue to provide capacity building support to local organisations. EPM conducts formal data collection on a bi-annual basis to understand the replication's effectiveness and sustainability and is in continuous communication with the brand organisations.

Across the whole ecosystem, EPM also mobilises new players and resources to advance the scaling up of impact (such as funders, topic experts and media). As an example, EPM collaborated with Tencent, Caixin Weekly Magazine, WeWork and Junlan Hotel Group to help 19 products on its platform successfully raise 26.5m RMB

(3.7m USD) in three days (through '99 charity day', one of the biggest online public fundraising events in China).

The case of Rainbow Village, a brand organisation that aims to support children whose parents are in prison, gives a more concrete picture of how EPM's infrastructure has supported interventions to scale.

Rainbow Village's product consists of scholarship support for education, facilitated visits to parents, summer camps, and psychological support, depending on the child's needs. Rainbow Village joined EPM in 2016 and, at that point, had seven local organisations replicating its model. By the end of 2019, this had grown to 60 local organisations, representing a compound annual growth rate of 79% for local organisation partners, and 59% compound annual growth rate (CAGR) for direct beneficiary growth (from supporting 368 children a year in 2016 to 2,691 children a year in 2019).

Rainbow Village: Speed of replication by number of direct beneficiaries and local replication partners



EPM has provided scale up grants of 410k RMB (57k USD) to Rainbow Village, and coordinated 10 roadshows with local hubs. Rainbow Village received training on how to turn its project into a product and support the fundraising of local organisations. EPM also connected Rainbow Village to other funders bringing in an additional 4m RMB (563k USD) to support their scale-up.

How EPM goes beyond traditional product platforms

Product platforms are not a new idea in themselves. Many initiatives exist to aggregate the most effective products or interventions in certain topic areas to help funders and practitioners know what to invest in. The What Works Clearinghouse, created by the US Department of Education, is a platform that aggregates the evidence base for interventions in education to make it easy for practitioners to identify and replicate proven interventions.

In most cities in the US, where local organisational capacity is less of an issue and where there is a mature and established market of non-profit organisations and funders, a product platform addressing the upstream challenge alone, may be enough to facilitate effective replication. However, in developing countries and some parts of developed countries, where the downstream players do not exist or lack capacity, and where local government or local corporate donations are still immature, we cannot focus only on the aggregation of the best products, we must also find a way to create the operating models that build local organisational capacity to mobilise resources and implement services sustainably.

EPM's model is distinct as an early effort to connect hubs and product platforms to strengthen the connection between upstream and downstream efforts for scaling. Hubs show the potential to play the essential role of building the capacity of local organisations both to raise resources from local government and local corporates and to develop their ability to implement the project. As we saw in the example in Guizhou, this dramatically raises the resources available for scaling. With hubs playing this role, the brand organisations don't need to be responsible for fundraising for their whole network of local organisations, where it will never be able to raise enough to scale to meet the size of the problem.

The other major benefit of using hubs is their neutrality. Part of the problem with existing approaches to scaling, especially when one large NGO receives major funding from a donor to scale its model, is that the local organisations it partners with become

dependent on it. Instead, when local hubs and organisations are neutral, they are empowered to work with multiple brand organisations in a marketplace approach, where local needs take precedence over donor priorities. This is usually a more sustainable model too as it allows hubs and local organisations to diversify their economics across a portfolio of different products.

Once a full network of mature hubs exists nationally, brand organisations and their funders do not need to source and manage a large network of local organisation partners by themselves; they can instead seek to 'plug and scale' into the network of neutral hubs and leverage the hubs' capacity, which gives access to the largely untapped market of local government and local corporate resources.

The challenges, limitations, and improvement areas for EPM in China

EPM's progress shows that there is potential in this approach to more quickly and sustainably move toward solving social problems at scale. However, there are still clear areas for improvement.

First, EPM estimates that half of the 38 local hubs are functioning well and the other half need more support from funders and incubators to build their aggregation capabilities. Only three provincial hubs are what EPM would consider 'high functioning,' and have proven an ability to tap into the resources of local governments and local corporates at a significant scale. Becoming a mature hub is challenging: it is highly dependent on local leadership, their drive, and the strength of their relationship network. EPM is doing further research on hubs to understand their development and how to help more hubs accelerate their development.

Second, the product platform needs to continue to push for higher quality products and more impact measurement. EPM has received criticism that it has prioritised "easy to replicate" products. This is a reflection of the Chinese social sector overall. Given the relative youth of the sector, there is widespread under-investment in monitoring,

evaluation and learning, with very few capable suppliers of impact measurement and action-research. As a result, most products/models do not yet have a clear evidence base for their effectiveness. As local capacity grows through the work of hubs, higher quality and more complex products can be replicated. This process could be accelerated if more donors in China were willing to fund professional impact measurement and action research into both intervention and operational delivery models.

Finally, there are still large policy gaps at national and local levels, so many social services lack the framework to mobilise the resources needed to reach the last mile. EPM is not sufficient to address these challenges alone. For example, a major reason Huiling's model has struggled to reach scale is that, while the government has established strong subsidies for children with IDD, adults with IDD incur a far greater cost per year, and have far less access to subsidy financing. Other organisations will need to step in to help address these policy gaps. In the US, it was the large parents' association (ARC) that drove policy advocacy from the 1950s to 1980s to expand access to services for adults with IDD in the USA.

Opportunities to build the infrastructure for scale in other countries: India as a case

The experience of EPM in China is relevant for other large countries looking to accelerate replication of social service development, and in some cases like India, there are examples where others have already taken these ideas further.

China and India are the two most populous countries on Earth, but there are major differences that shape execution. First, in contrast to China, India has a much more mature civil society and philanthropic landscape, with many large brand organisations that have reached nation-wide scale, and many more mature foundations engaging in strategic grant-making and philanthropy. Second, the government plays a much bigger and more centralised role in China, whereas India is much more decentralised, with clearer distinctions between national and state governments. In China, the government

provides the major funding directly for social services, whereas in India, the government taps CSR as a means to fund social services; their law requires that 2% of all corporate revenues be spent on CSR activities.

For upstream aggregation, given the large numbers of brand organisations and relative maturity of the market, India could look to establishing 'issue specific' product-platforms. For example, it is not uncommon to find non-profits working in 'skill development' with similar models, with varying levels of evidence on how effective the models are. Issue specific product-platforms could help aggregate different models and profile which models work best for different segments, industries, or geographies.

For downstream aggregation, India could look to establish collaborative networks of neutral hub organisations that could help to capacity build and mobilise local resources for grassroots civil society organisations. These hubs could represent certain beneficiary segments or populations, which face similar challenges by geography.

A great example is the issue of internal migrant workers in India. India has approximately 140m internal migrants, that seasonally migrate in the agricultural offseason, to find work in cities, often in construction and manufacturing. Migrant workers face low wages, poor working conditions, and lack of access to social security; and this situation has been made much worse, post COVID-19. 90% of migrant workers lost their livelihoods within 3 weeks of the April lock down, with approximately more than 80% lacking access to basic government entitlements and social security⁴.

Migrant workers are disproportionately made up of excluded castes and discriminated communities. Supporting these workers to access social security and safe livelihoods requires the intervention model, but more importantly, requires a delivery model by local community-based organisations that are intimately familiar with the geography and challenges faced by these specific communities.

⁴ Voices of the Invisible Citizens Report – Jan Sahas Survey, (n=3,196, April 2020)

Jan Sahas is a leading NGO in India supporting migrant workers, and has long recognised the need for a grassroots-representing-grassroots approach. Post COVID-19, Jan Sahas launched and established the "Migrants Resilience Collaborative" (MRC), India's largest grassroots-led collaborative, and largest non-governmental initiative dedicated to migrant workers and their families. MRC aims to support 10 million workers and their families in 100 districts and cities (across source and destination) over the next 5 years. In the first few months, MRC has already set up a network of 72 District Migration Resource Centers (DMRC) to provide assistance and support to migrant workers in 11 key states across India.

Importantly, these district migration resource centers are neutral and not tied to any one NGO but instead, work with multiple civil society organisations across the MRC network to aggregate interventions needed for local migrant populations. MRC will continue to aggregate and improve the 'upstream' interventions needed for migrant workers: access to social security, access to safe livelihoods/responsible recruitment, financial inclusion and worker protection, and will facilitate the replication of the 'downstream' delivery models through its network of resource centres and partners.

There are potential learnings from both countries. EPM shows the value of aggregating products and geographic/community focused, neutral hubs (instead of each NGO developing their own network), and India shows the value of applying the same thinking with a deep issue-based lens e.g. migrant workers to build the depth, as well as the breadth of impact.

Conclusion

In emerging social sectors, the single organisation led approach to replication and scaling impact is not enough. Aggregators are needed to build the upstream (product platforms) and downstream (hubs) infrastructure to accelerate scaling. Once built, single organisation or model led approaches can effectively use the infrastructure to "plug and scale", instead of having to work through all the upstream and downstream activities for scaling by themselves.

For scale-up and systems change funders, investing in the infrastructure for scaling is critical to maximizing their return on impact. In emerging social sectors, scale-up funding could be wasted if the downstream infrastructure does not exist to sustain and scale replication. Investing in this infrastructure helps scale and replicate the best models for local needs, and can exponentially leverage the resources of local governments and corporates, so that we can effectively address the scale of the social problems we seek to solve.

Warren is the founder and Managing Director of GDI's East Asia office, leading GDI teams across mainland China, Hong Kong SAR, and India. Global Development Incubator (GDI) is an incubator for development ventures that unlock transformative social impact at scale. GDI has incubated more than 30 collective impact initiatives and funds; many are multi-million USD efforts tackling systems change. For us, scalability and impact aren't just a metric of success, they're criteria for support. We make big bets that can push the limits of our current systems, or outright transform them. Warren can be reached at warren.ang @globaldevincubator.org

Yanni Peng is CEO of Narada Foundation. Narada Foundation, founded in May 2007, is a private nationwide foundation approved by the Ministry of Civil Affairs of China. Narada Foundation is dedicated to the mission of fostering China's third sector. It has been proactively shaping a healthy ecosystem and providing public goods for China's philanthropic sector ever since. In the service chain of the sector, Narada Foundation provides funding and resources to support the development of effective philanthropy projects and non-profit organisations and to catalyse grassroots social innovations. Yanni can be reached at pengyanni@nandu.org.cn.