

ORGANIZATIONAL MECHANISMS OF INCLUSIVE GROWTH: A CRITICAL REALIST PERSPECTIVE ON SCALING

Christian Seelos

Johanna Mair

ORGANIZATIONAL MECHANISMS OF INCLUSIVE GROWTH: A CRITICAL REALIST PERSPECTIVE ON SCALING

Christian Seelos¹

Johanna Mair¹

Abstract

We investigate the challenge how the outcomes of innovation for inclusive growth, the novel organizational recipes, can be scaled to match the dimension of poverty. We conceptualize scaling as sustained event regularities between doing A and expected outcomes B. Building on a critical realist perspective, we develop an analytical framework of organizational closure and apply it to an extreme case, an organization with an inclusive growth model that has sustained event regularities for more than two decades. Our analysis reveals closure as an organizational competence with important implications for achieving scale in the context of poverty. We develop a number of propositions between the link of organizational closure and scaling with implications for practice and further academic research.

Keywords: organizational closure, social entrepreneurship, counterfactual analysis, retrodution.

Acknowledgements: We thank Funda Sezgi for excellent support with field interviews and qualitative analysis of data. We are most grateful to Aravind staff and in particular to Ravilla Thulasiraj for their generosity and inspiration.

¹ Professor of Strategic Management, IESE

ORGANIZATIONAL MECHANISMS OF INCLUSIVE GROWTH: A CRITICAL REALIST PERSPECTIVE ON SCALING

1. Introduction

In gastronomy as in economic development, scaling up a good recipe serves the purpose of feeding more people. However, while scaling in gastronomy seems unproblematic this is not so in economic development. One important recipe many thought scalable was the set of policy prescriptions for economic growth formulated under the framework of the Washington Consensus 20 years ago. But decades of investments of significant economic and development resources have “scaled” 80% of global wealth into the hands of less than 10% of the world population (Davies, Sandström, Shorrocks and Wolff, 2008). Poverty persists in important and intolerable dimensions (Chen and Ravallion, 2007; Easterly, 2002). What makes scaling recipes for economic development so difficult?

Some have singled out inadequate understanding of cause/effect relations or simply the application of the wrong development recipes as the perpetrator. In 2005, the World Bank published a report titled “Economic Growth in the 1990s: Learning from a Decade of Reform.” Rodrik (2006, p. 974) calls it “...a rather extraordinary document insofar as it shows how far we have come from the original Washington Consensus. There are no confident assertions here of what works and what doesn’t – and no blueprints for policy makers to adopt. The emphasis is on the need for humility, for policy diversity, for selective and modest reforms, and for experimentation.” Others focus on the inadequacy of context for a given recipe. For example, a recent report by the Commission on Growth and Development (2008, p. 4) states: “Economists know how markets work, and they can say with some confidence how a mature market economy will respond to their policy prescriptions. But mature markets rely on deep institutional underpinnings, institutions that define property rights, enforce contracts, convey prices, and bridge informational gaps between buyers and sellers. Developing countries often lack these market and regulatory institutions. [...] The impact of policy shifts and reforms is therefore harder to predict accurately in a developing economy.”

An emerging development paradigm that focuses on inclusive growth through the integration of the poor now bases its hopes on bottom up, entrepreneurial, innovative solutions. Some authors claim that so-called social entrepreneurs create novel recipes in the form of innovative business models able to serve the poor efficiently (Seelos and Mair, 2005). The hopes for this new paradigm are such that one prominent social entrepreneur was recently awarded the Nobel Peace Prize. With all the focus on innovation, we wish to draw attention to the challenge of

how to take the outcomes of innovation, i.e. the novel recipes, and make them bigger or replicate them; in other words: how to achieve scale. The problems faced by small innovative and entrepreneurial organizations vis-à-vis the dimensions of poverty are a fundamental challenge for inclusive growth. *“It is often recognised that most of these grassroots initiatives are small, underfunded, poorly staffed, slow and localised in the face of poverty, hunger and degradation on a vast scale. They are thus considered to be only actions at the margins, capable of providing local relief and empowerment, but not of tackling the real issues of the eradication of hunger and poverty for hundreds of millions of people”*. (Uvin, 1995, p. 927). Bradach (2003) has similarly expressed concerns about the fragmentation, smallness and insufficient realization of the potential of proven programs.

Scaling as used in this paper implies an understanding of relevant cause/effect relationships. It constitutes a bet on the ability to generate event regularities, i.e. how doing A produces more of outcome B. We investigate scaling from the perspective of a critical realist philosophy of science by following Sayer (1992, p. 122) in understanding that *“Realist philosophy reflects upon the conditions which must hold if regularities are actually to occur, that is, it asks what a system and its constituent objects must be like for regularities to be produced.”* Guiding our inquiry for this paper is the principal question: What does an organization need to be like to produce the event regularities required for scaling? This shifts the focus from the adequacy of recipes or the adequacy of context to the particular organizational characteristics in which we expect recipes to create desired and expected outcomes.

We develop our analysis as follows. First, we reflect on the implications of applying a critical realist philosophy of science to the challenge of scaling. This allows us to create an analytical framework that characterizes the “anatomy and physiology” of organizations as experimental production platforms able to generate expected outcomes on a sustained basis. We use rich and original data on an extreme case, an Indian eye hospital, the largest in the world, to illustrate our analytical model and to develop a new category of organizational competencies: closure competencies. Finally, we develop a series of propositions for scaling mechanisms relevant for inclusive growth in underdeveloped market contexts. We end by providing implications of our findings for research and practice.

We wish to make several contributions to the nascent literature around scaling in the context of inclusive models for growth. First, we develop a concept of scaling in an empirically unambiguous manner adequate for scholarly investigation. Second we operate from an explicit and transparent commitment to ontology and epistemology and shape our analytical design accordingly. This allows us better to compare and contrast our analysis with relevant findings by other scholars. Finally, we wish to reconnect to the early organization literature with its assumptions and perspectives about closed systems, much of which was a byproduct of the search for improved efficiency or performance as Thompson (1967) pointed out.

Critical Realist Reflections on Scaling

Contemporary perspectives on organizations predominantly employ an open systems view that is based on multiple internal and external relationships and dependencies among organizations and their environments (Baum and Rowley, 2005). Critical realist scholars also maintain that closure in social systems is not possible (Archer, 1995; Bhaskar, 1975; Sayer, 1992). The open systems nature of social phenomena unfortunately limits the kind of predictive success implicit in a

theory of scaling, i.e. that doing A will create outcome B. As Sayer (1992, p. 130) states, the notion of predictive success as the primary goal of any science, natural or social (...) “*is certainly false as regards social science and probably false for many natural sciences. In comparison with the number of explanations produced, predictions are relatively rare, especially accurate ones.*”

To the contrary, we believe that a perspective on closure is relevant and insightful for understanding the challenges to scaling of inclusive organizational models particularly in a context of mass poverty. The traditional closed system literatures on scientific and administrative management or bureaucracies primarily focused on formal structures and internal variables (Thompson, 1967). In our investigations, we engage a critical realist perspective of closure that focuses on the coupling between structures, social mechanisms and outcomes to provide an explanation of how organizational closure is possible and how this relates to explicit and measurable dimensions of scaling.

How Critical Realist Perspectives Inform This Study

Critical realism as a philosophy of science was originally developed and substantially commented on by scholars including Harré and Madden (1975); Bhaskar (1975); Collier (1994); Archer (1995); Sayer (1992); Lawson (1997), and Bunge (2006). Realist perspectives were also proposed as fruitful avenues for investigation in organization and management studies (Ackroyd and Fleetwood, 2000; Durand and Vaara, 2010; Fairclough, 2005; Johnson and Duberley, 2000; Reed, 2005; Tsang and Kwan, 1999; Tsoukas, 1989; Van de Ven, 2007, and Whitley, 1984). A hallmark of critical realism is the ontological commitment that differentiates reality into three dimensions. The *empirical dimension* refers to empirically observable events. They constitute the outcomes caused by a number of mechanisms operating in parallel. The link between mechanisms and empirical outcomes represents the second dimension of the critical realist ontology, the *actual dimension*. Mechanisms are enabled by structural configurations and thus by the powers of objects and their structural relations with each other. This is the third dimension, the *real dimension*. It specifies the link between powers, mechanisms and empirically observable events. If the objects under investigation are humans, then the term “powers” refers to the set of ways in which people can act (manifesting as mechanisms in operation) when stimulated externally or intrinsically. Thus, while the power to run is possessed by all humans who have legs, the power to fly is not. Powers also reflect social structures. For example the hierarchical social structure of manager and direct report engenders in the manager a power to fire the direct report but not the other way around.

A commitment to this differentiated ontology necessitates the view that – outside of strictly controlled experimental systems, i.e. in the real world – there are multiple possible correlations between a cause and an effect. The reason is that the three ontological layers are irreducible to each other and are out of phase with one another (Lawson, 1995). Observable outcomes are out of phase with mechanisms because many mechanisms operate in parallel and may reinforce each other in unexpected ways or cancel each other out. Mechanisms may operate below a critical threshold and thus their effects remain unrealized. Mechanisms are also out of phase with the structures that enable them. They may not be adequately triggered or self-conscious actors may choose to act in different ways in the same situational context. Thus, realist explanation requires that “Science is the illumination and elaboration of the structures and mechanisms that govern the events of experience” (Lawson, 1995, p. 13). The consequence is that regular realizations of causes and effects, i.e. event regularities, are not expected outside strictly controlled experimental systems. The best we can expect are what Lawson (1995) calls

“demi-regs”: temporary event regularities that create short-lived and spatially restricted patterns: doing A may sometimes create outcome B or doing A has a tendency in certain situations to create outcome B.

The concept of scaling implies two important dimensions of event regularities. First, it requires the existence of some concrete object that is scalable, e.g. an organization that regularly produces a desired outcome such as a particular type of product or service. If no regular input-output mechanisms are in place, the concept of scaling tends to refer to a discourse where any input or output can be framed as indicative of scaling, e.g. running more short-term projects may constitute an instance of scaling. However, this gives rise to an “anything goes” epistemology where truth is just a discourse. Secondly, scaling constitutes a bet that doing more of A or doing A better is causally linked to create more of a desired outcome B. Implicit in this perspective on scaling is a notion of temporality. While short-lived event regularities are unproblematic, we argue that scaling in the context of inclusive growth and global poverty requires sustained event regularities over significant spans of time to match the dimensions of needs.

The critical realist perspective, while highlighting the indeterminacy of real world open systems, also presents an opportunity to escape this indeterminacy. The literature is ambiguous about whether and how it is possible to create closure in the sense of a production system for repeated event regularities. While Tsoukas (1989) postulates the impossibility of constructing closure conditions in the social sciences, the same author considers the opportunity to escape the indeterminacy of open systems: *“In other words, management must create conditions of organisational quasi-closure so that certain activities of interest are controlled [...] and particular results are obtained. Thus although the causal powers of management operate in open systems it is only when quasi-closed systems are constructed that a set of desirable regularities accrues”*. (Tsoukas, 2000, p. 40).

Methods and Analytical Framework

Methods and Data

Our objective is to explain regularities in achieving predetermined organizational objectives. Following a critical realist research tradition we focus on identifying causal mechanisms and how they work, and on discovering if they have been activated and under what conditions (Sayer, 1992). A logic of retroduction takes us “behind the surface phenomenon to its causes, or more generally from phenomena lying at one level to causes often lying at a different deeper one” (Lawson, 2003, p. 80). This is achieved by following prescriptions of critical realist scholars regarding research design and analysis. We develop an analytical framework in the form of a theoretical model that links structures and their causal powers, mechanisms and outcomes to derive a causal explanation of regularities in organizational outcomes. We then scrutinize the model empirically (Leca and Naccache, 2006). We do so by basing our analysis on one in-depth case study, creatively using alternative data sources, emphasising processes, and identifying mechanisms *“by continuing to ask the question why and collecting more data until we believe we have an explanation”* (Easton, 2000, p. 217).

We chose to study Aravind as an instrumental and extreme case to understand what an organization needs to be like in order to produce the outcome regularities required for scaling. Scholars have provided detailed descriptions of Aravind’s ability to consistently produce

tangible and clearly measurable outcomes over many years (De Véricourt and Sousa Lobo, 2009; Rangan and Thulasiraj, 2007). Our analysis benefitted from superior quality of data due to the privileged access of the research team to the organization over an extended period of time (from October 2004 until November 2009).

When gathering and analyzing data, critical realist researchers use discourse analysis but also go beyond in order to reach the domain of the actual. Central to this approach is that researchers take a critical stance (Leca and Naccache, 2006) and observe actors' actions and practices beyond the discourses they develop (Archer, 2002). We have conducted multiple fieldtrips to India and triangulated data from more than 51 interviews by interviewing: 1) the same person several times and in different (group or geographical) settings,¹ 2) a number of informants on the same practices and events, and 3) experts, external collaborators and members of support organizations. We corroborated interview data with reports and studies on the focal organization conducted by independent researchers, direct observations of practices by four researchers at different times, participation of two researchers in management meetings and retreats of the management team, archival data including internal reports and presentations including performance data on the various hospitals, and physical artefacts. Appendix A provides a summary of interview data.

Analytical Framework

Sayer (1992) points to the need to specify the particular characteristics of a system and its constituents that would explain the generation of event regularities. Bunge (2006) defines the minimal required set of specifications for a material system (such as an organization) as: "constituents", "structure", "mechanisms" and "environments." We apply this set of four system variables to the study of organizations by making the following adjustments:

1. We limit "constituents" to human actors; all non-human constituents such as machinery are explanatorily unproblematic for the purpose of this study because they do not create systemic unpredictable variance that undermines the creation of event regularities.
2. "Structure", following Hodgson (2007) and Tsoukas (1989), refers to the set of human relations that have both enabling and constraining effects on the generation of mechanisms.
3. Because power resides in the structures of objects and their relations, the variables "actors" and "structure" together determine the system potential and the feasible set of mechanisms that can be generated (Harre and Madden, 1975; Bhaskar, 1975).
4. "Mechanisms" refer to the set of distinct processes that make a system "*what it is and the peculiar ways it changes*" (Bunge, 2006, p. 126). Mechanisms in this sense are also semantically equal to organizational "events" and thus refer to both causes as well as effects.
5. Because organizations are neither isolated nor independent of their environments, it is the set of organization internal and environmental powers and operative mechanisms that determine the outcomes an organization can achieve.

¹ For instance, we interviewed the Executive Director seven times during the course of our study. The interviews were conducted by different researchers, at different occasions and in different locations/continents.

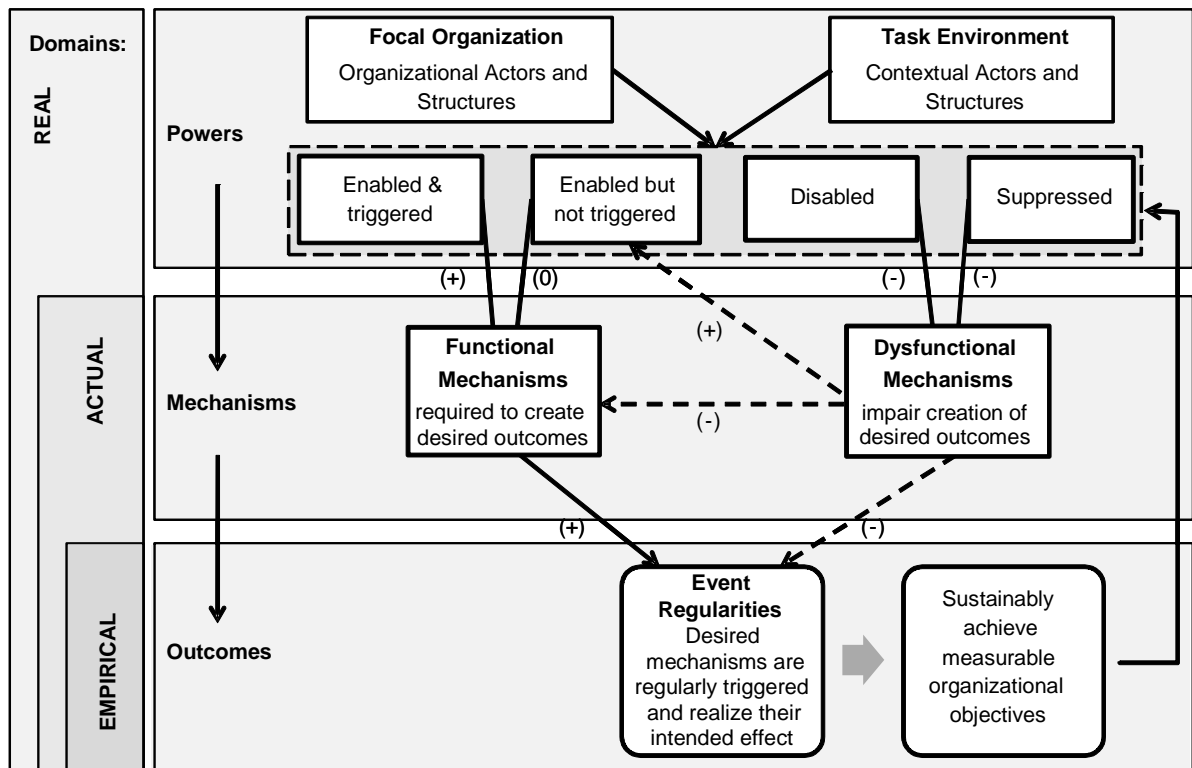
6. We treat the variable "environment" as the relevant actors, structures and mechanisms in the task environment of the focal organization.

While actors, structure, mechanisms, and environment specify the content of our framework, we need to integrate the specific coupling between these variables and outcomes. We map the above variables to the differentiated ontology of critical realism comprised of the real, actual and empirical domains as depicted in Figure 1. Since we are explaining organizational closure, our framework needs to specify the organizational conditions for event regularities. We treat the empirical observation of consistent achievements of concrete organizational objectives as the explanandum, i.e. the event regularity, and actors, structures and mechanisms as the explanans. However, the objectives are themselves to some extent also explanatory of the particular organizational set-up. This is highlighted by the arrow in Figure 1 that links the empirical domain with the deeper structure of powers residing in the real domain. Objectives shape decisions about which mechanisms to enable, trigger, disable or suppress. Furthermore, maintaining a set of concrete objectives over time may be a prerequisite for sustaining organizational closure as further detailed in our illustrative case study.

A central aspect of our analytical framework is Harré and Madden's (1975) notion of enabled mechanisms. To the extent that functional mechanisms are part of a causal chain of processes and events necessary to achieve the organization's objectives they need to be enabled. "Enabled" means that they are part of the set of properties that define an object, i.e. properties that reside in system-relevant actors and/or their relations. For example, doctors in an eye hospital need to have training in cataract surgery if providing this service is an organizational objective. The functional mechanism of doing eye surgery is enabled through training or through hiring trained eye surgeons. However, enabled functional mechanisms constitute a potentiality that may not be realized. As Bhaskar (1975, p. 43) explains: *"The experimental scientist must perform two essential functions in an experiment. First, he must trigger the mechanism under study to ensure that it is active; and secondly, he must prevent any interference with the operation of the mechanism. These activities could be designated as 'experimental production' and 'experimental control'."* In our example, the enabled functional mechanism of surgeries also needs to be triggered, for example by the presence of patients, the availability of proper equipment and by an adequate incentive system that motivates the doctor to do surgeries. Furthermore, experimental production and control as the hallmark of organizational closure also require that negatively interfering mechanisms, i.e. dysfunctional mechanisms, are disabled or suppressed. Otherwise enabled functional mechanisms may not be triggered, or if triggered may be neutralized or dominated by dysfunctional mechanisms, and outcome regularities cannot be achieved on a sustained basis as indicated by the dotted lines in Figure 1. Thus, the achievement of organizational closure requires managerial effort in the same way that *"...the experimentalist's task is to manipulate the entire experimental system, so as to manufacture the desired interrelationship between independent and dependent variable. The experimentalist is indeed a system builder and the crucial evidence is produced not by controlled observation but by work."* (Pawson and Tilley, 1997, p. 60). We thus conceptualize closure as the prerequisite for sustained outcome regularities in our framework. It requires that functional mechanisms, i.e. mechanisms necessary to achieve organizational objectives as their effect, are enabled and triggered on a sustained basis and that dysfunctional mechanisms, i.e. those that would prevent the realization of the effects of functional mechanisms, are disabled or suppressed. We summarize our analytical framework in Figure 1.

Figure 1

Analytical Framework – Organizational closure as requirement for the sustained achievement of concrete objectives



Note: The signs (+), (-), and (0) refer to positive, negative or no effects, respectively.

Illustrative case study

We use an illustrative case study for two main purposes. First, it allows us to bring the analytical framework alive and to demonstrate its usefulness. In particular, we wish to emphasize the fruitfulness of engaging in counterfactual arguments. The analytical framework makes it clear that empirical events are also “caused” by the absence of dysfunctional mechanisms. Why are they not happening? Since open systems, not closure, are the norm, we need to ask why expected variance is not realized but why we see sustained patterns of regular events. Postulating counterfactuals for realist organizational analysis thus involves asking questions like: “what could or should have happened but did not”? It is only when functional mechanisms are sustainably triggered and are undisturbed by dysfunctional mechanisms (or when the effect of dysfunctional mechanisms can be understood and controlled) that closure can be achieved and regular events that lead to expected outcomes can be realized on a sustained basis. Our second reason for using this case study is that the focal organization, Aravind, has been engaged in a number of scaling mechanisms over a long time. Iterating between the analytical framework and Aravind’s rich empirical history enables us to create and support a number of propositions on scaling mechanisms formulated below.

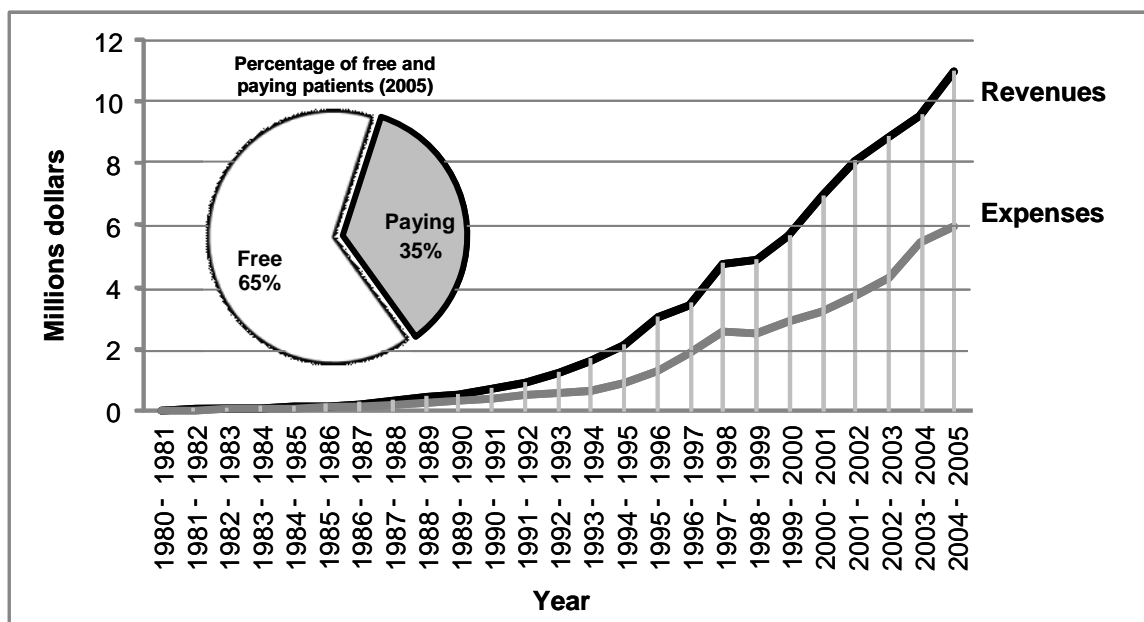
Aravind Eye Hospital

Aravind in India is the largest group of eye hospitals in the world. The group's mission is to eradicate unnecessary blindness and it has pioneered a novel approach to delivering eye surgery for cataracts that integrates free surgery for the poor as a major strategic objective. Cataracts affect millions of people all over the world and are routinely treated in wealthy countries by replacing the clouded natural lens of the eye with an artificial lens to restore sight. However, the market prices of lenses and surgery have traditionally been out of the reach of poor people. Partial or full blindness limits or prevents the ability of the poor to engage in economic activity. In addition to the personal hardships of the blind in poor countries, limited social welfare and insurance systems put a severe burden on the supporting families and their small private savings causing significant economic losses and social and individual suffering. According to the World Health Organization, the economic costs of blindness are estimated at United States \$25 billion. Access to quality eye care and appropriate technologies would make up to 80% of blindness preventable or curable (World Health Organization, 1997; World Health Organization, 1999).

In 1976, Dr. Govindappa Venkataswamy ("Dr. V" as he is called at Aravind), a retired ophthalmologist, founded Aravind in the city of Madurai in the Indian state of Tamil Nadu. Over the years Aravind has expanded to a group of five hospitals that currently perform over 300,000 eye surgeries annually and provide eye care services to more than two million outpatients. While more than half of the high-quality eye surgeries are provided to the poorest for free, Aravind manages to generate significant operating profits (see Figure 2). Profits are used to invest in capacity building and increasing the scope and scale of Aravind's activities. This sustained growth of the delivery of a highly standardized service over two decades constitutes an event regularity that requires closure as an explanation from the perspective of critical realism.

Figure 2

Revenues and expenses of Aravind between 1980 and 2005 and percentage of free versus paying patients treated in 2005



Note: Data provided by R. D. Thulasiraj, Executive Director of the Lions Aravind Institute of Community Ophthalmology.

How Aravind achieves closure

The mission of Aravind is to eradicate needless blindness. Three core objectives define Aravind's strategy: 1) providing high volumes of high quality cataract surgery; 2) maintaining a high ratio of poor patients to paying patients (70:30), and 3) profitability. Here, we highlight a number of key functional mechanisms required to achieve Aravind's mission and the ways in which they are enabled and triggered. At the same time we emphasize how potential dysfunctional mechanisms are disabled or suppressed.

Functional Mechanism (1) – maintaining focus on core services over sustained periods of time

From the beginning Aravind has focused on cataract treatment as its prime service. The focus on cataracts is relevant given the mission of eradicating needless blindness as an Aravind doctor confirms: *“Blindness is growing. About 330,000 every year is [the] incidence of cataracts in this area. [The number of] surgeries done in this area was about 75,000. Every year, a backlog of about 225,000 more blind people is pulling the society back”* (Doctor). However, the many needs in a context of large-scale poverty may tempt organizations to expand their scope and thus seek to “scale” their activities across many dimensions of needs. But this adds complexity and contributes to loss of control over too wide a range of organizational and contextual variables, which, in turn, constitutes a loss of closure. Aravind chooses to maintain its original goal of eradicating needless blindness and cataract surgery is perhaps the best mechanism to achieve that. Thus, two thirds of all eye surgeries performed by Aravind are cataract surgeries, while the other third combines another eleven different types of eye surgeries.

Before his death, Aravind's founder, Dr. V, encouraged a number of relatives, many of whom are eye doctors, to work within the organization. Today they keep Aravind's dedication to its mission alive. To further suppress tendencies for “mission drift,” Aravind's board has recently agreed on a new “stretch goal:” to build the capacity to provide one million cataract surgeries annually by 2012. The commitment to this goal suppresses dysfunctional mechanisms such as doctors pursuing their pet projects or investing too broadly into the provision of specialty ophthalmology services (unless these are necessary to the core mission as detailed below).

Functional Mechanism (2) – maintain high quality of medical services for all

Aravind provides the same quality treatment to both paying and non-paying patients. This policy disables or suppresses a number of dysfunctional mechanisms. Having only one category of patients eliminates ambiguity for nurses or doctors about what type of surgical procedure would be adequate for a given patient. Misinterpretation and conflict is disabled and unnecessary complexity is removed from the system. Staying true to this mission also suppresses feelings of distrust or cynicism among employees which would counter Aravind's need for high employee dedication and motivation: *“Every employee is very proud of us. Even a gardener [...]. When they feel pride, they feel without them it's not going to function. When there is this right [culture], I think so many things follow”* (Doctor). *“[If the] patient is satisfied with me, I am happy too. Without these things, we can't develop our hospital. Hard work is needed here, we [all] contribute”* (Nurse). The commitment to serving all patients regardless of income or background and systematically suppressing any tendencies to prioritize paying over non-paying patients (for example, to increase profits) are crucial mechanisms for Aravind: *“Why is Aravind unique? We have a huge patient load. We have so much technology. We have 250 doctors. We have all. But as I said, we are modest. We don't cheat our patients. We are not greedy for money. We always do the best for the patient. We always respect people”* (Founding Member).

Functional Mechanism (3) – provide high volume of services through sustained access to key resources

Because of the high incidence of blindness, achieving high-volumes of cataract surgeries is essential to the success of Aravind's mission. We elaborate on the relation between our analytical framework of organizational closure and scaling in the next section. Here we focus on the requirement of sustaining access to the right resources as one prerequisite for achieving a high volume of services (increased productivity as an additional prerequisite is discussed below). Because our analysis is geared towards explaining closure, the resources of interest here are primarily human resources such as doctors, nurses and family members. It follows from our analytical framework that to reach an adequate degree of closure and outcome regularities, Aravind needs to ensure access to actors with powers to provide high volumes of high quality services and the sensitivity to be triggered into action on a sustained basis. The repertoire of mechanisms that Aravind deploys includes the building of reputation and trust in the task environment, the internalization of training that emphasizes skills and values, and establishing a rigorous selection process for various actors.

The reputation and trust that Aravind has been able to build over the last 30 years has enabled the organization to ensure a continuous supply of nurses and doctors. Recruiting girls from rural India and training them as professional nurses required Aravind to break with dysfunctional norms and traditions. For example, girls are supposed to get married at the age of fourteen and are not encouraged to leave their villages to work in cities such as Madurai. *“Also it is very difficult to recruit girls in this part: here, the psychology is that there are social arrangements [marriage] and all. Families were afraid to send their daughters. Now they're comfortable with it. Yes, that was one of the advantages of Aravind in the south area: they knew the organization. If not to other organizations, they'd send to Aravind. [...] Now, we're getting lots of applications”* (Hospital Administrator). It took Aravind several years to gain legitimacy as a trustworthy institution: *“The families of those girls, they feel that the girls are safe, they are serving for a good cause, they have values of [the] culture of Aravind, and they're serving the people. So, they are very happy. And first year, there is a girl from the family, next year relatives are here, because it's good here”* (Chairman).

While legitimacy and reputation ensure the supply of resources/actors, it is the content of the training, i.e., a combination of skill focused and values based training that generates the necessary attitudes to trigger the right functional mechanisms within Aravind's system. *“Training is bi-directional. Our paramedics and doctors have to be professional and very efficient and competent. [The next most important thing is having] ... ethics equal to the values of Aravind. How to perform beyond 100%. Hard work takes you to 98%, knowledge takes you to 96%. But attitude will take you to 100%”* (Chief Medical Officer).

In order to generate regular and predictable outcomes training programs need to be consistent. A possibly dysfunctional mechanism therefore lies in the fluctuation of training personnel. Aravind disables such tendencies by the direct involvement of family members in both skills and values training. The following description underlines this process whereby family members become the carriers of skills and values and their involvement disables dysfunctional mechanisms: *“Basically first of all the training that I received at Aravind. ... I have observed many surgeries. I have observed different techniques, machines, all the aspects to run a hospital. ... The next thing I was lucky enough to meet Dr V. Always he emphasized on the attitude and knowledge adaptation. [...] Other organizations, I don't know how they give emphasis on the vision and mission. Maybe they*

are on [...] paper, but in Aravind it's not like that, it is not a piece of paper. It is actually taught, trained and they ask us to perform it" (Hospital Administrator).

As important as the quality of training to achieve outcome regularities is the decision of who is hired, as well as how disciplined and aligned with the culture the individual actor – nurse, administrator or doctor – is in performing her tasks. In order to disable potential dysfunctional mechanisms or avoid hiring the “wrong people” Aravind engages in a meticulous selection process: *“When we select a person, both consciously and unconsciously, the most important criteria is organizational fit. At every level. And the more senior they are, the more rigorous is the assessment process. Like, for example, taking a doctor, because a doctor's position by default is an influential position in the organization: people listen to them, the nurses and the people. If they set wrong standards, that would kind of dilute the organization, so we put a lot of emphasis there. We also want those candidates to feel comfortable in this work environment. [...] in selecting senior people, doctors and all that, we would often have them spend about three or four days with us before we make a decision” (Executive Director).*

A widely used mechanism to suppress behavior that is not aligned with Aravind's important cultural elements – compassion, transparency and integrity – is deemphasizing the reliance on individuals: *“We trust the common man rather than a VIP. We do not go for titles and positions. I can say to everyone, ‘I studied in Harvard 40 years back’. Who cares; it is all about my behavior. You will never see our titles anywhere: for the patients, I am Dr Natchiar, that is all. Only when they look at me, they think I am a senior, I am an old lady” (Founding Member).*

Reflecting on an instance where Aravind had to intervene to discipline a doctor, the executive director explains: *“Very, very rarely we've asked people to go. Maybe in the 30, 35 years, I can only think of maybe one or two instances. Very rare [...] Like even recently, I had to counsel an arrhythmia specialist, because he was not becoming a team player. He likes to publish. That is good. We appreciate that. And he is a good doctor. But then there was a time when he went overboard, like throughout the day he would want to do internet browsing. There are patients waiting. If internally they would tell something at the department head level, he would kind of do it but not with the spirit. Then I had to sit and chat with him, really making him understand the biggest loser is him, not us. Ultimately, it is you who is wasting time, and the biggest impact will be you, because this institution will go on after you leave.”*

Habit formation to ensure disciplined and uniform behavior also spans across all task domains as an important way to enable and trigger functional mechanisms without much conscious consideration, and to disable dysfunctional mechanisms: *“Northern girls are very loud and expressive compared to the southern Indian women. But in the hospital setting, you cannot afford being very loud. Everybody knows that operating skill is a skill that if you repeatedly do it, you will get it. But what is important is not the skill, value and attitude, that makes the difference. [...] My first job is to make them de-learn what they have learned during their undergraduation. After de-learning then we inculcate the Aravind poison” (Doctor).*

Functional Mechanism (4) – maintain focus on treating 70% of all patients for free

The decision to keep the ratio of non-paying to paying patients at 70:30 is highly relevant given Aravind's mission. Cataract services are available to patients who can afford the routine procedure but are out of reach of the poor. However, being poor and blind is a dramatic burden in developing countries because of the lack of welfare or insurance systems. Therefore, the highest need for cataract treatment resides in the poorest part of the population. In contrast to popular views, the needs of the poor do not automatically constitute a real demand. Instead,

Aravind needs proactively to generate demand. A number of dysfunctional mechanisms in Aravind's task environment need to be disabled or suppressed. Many poor people consider blindness an age related or otherwise normal event so the idea that this can be treated does not naturally occur to them. Many are also distrustful towards doctors and perceive a visit to a hospital as a disruption to their regular day to day lives. A further complication lies in the necessity of patient compliance, which normally includes repeated visits to a hospital for diagnosis, surgery and post-surgery monitoring. Lack of compliance is a critical dysfunctional mechanism preventing the generation of expected outcomes. Aravind enables a sustained level of large-volume demand generation from poor patients by organizing eye camps in rural villages. Because success depends on suppressing distrust, Aravind partners with local community groups or local politicians that have the trust and respect of the rural poor. Aravind has built up a large network of community partners to generate a robust structure for enabling and triggering camps on a routine basis, thus generating sufficient demand for its high-volume growth strategy. Non-compliance is disabled by taking patients through the whole process in the shortest possible timeframe: *"We have a system which would make sure that we would try to see off-station patients quickly. [...] If it is eye-camp, they really want to complete the loop in one visit, the whole service loop. If they want glasses, we give the glasses right away. If they need surgery, we have buses waiting and then taking them back, bringing them back, because there's too much hassle to make people come again and again"* (Executive Director).

This is also why the commitment to high quality is so important: *"As a policy we don't do any marketing, but what we do is, we do quality: our quality speaks. It's the word-of-mouth that helps us. [...] When you have your surgery done, when you have your satisfied patients that is your marketing tool. They go to the community and spread their satisfaction. It's not only the surgery it's the quality of services we provide"* (Hospital Administrator).

A further potential dysfunctional mechanism threatening closure could be the need for doctors to earn extra money from treating more paying patients, thus neglecting treatment of non-paying patients to the detriment of the 70:30 ratio. To disable this potential, Aravind functions on a strict no appointment basis. Patients receive treatment from Aravind, the organization, not any particular doctor: *"In [...] India, the doctor starts practice, becomes a very famous one, and starts a hospital. He's a single person; if he goes somewhere, the hospital is closed. There is no system if [the] hospital is based on one person only. What we are trying to do: people are going to the hospital, thinking that they'll be treated the best. [...] Other hospitals are doctor based: if a patient came to see me, next time, he should come to me again. But in a system like Aravind, it's not necessary. Today, I'm here, next time maybe not. Whoever is there will take care of him. If he wants to see a doctor, he has to come here in a particular date and wait for him. Because none of the Aravind doctors work [on an] appointment basis, we take whoever comes. This is the main difference from other hospitals: otherwise, some doctors may not get any patients while [an]other doctor is overloaded"* (Chief Medical Officer). In addition the constant presence of Aravind family members further suppresses any deviating tendencies, for example through fatigue from the daily routine or loss of motivation. *"Now what we are constantly telling them, that we should be sincere, that we should have passionate care, that 70% should be free and 30% paying and we should respect the patient. All those things you have to take them along and if they lose it everything is lost. They will not be able to run the institution successfully. Today Aravind is working not because of technology"* (Founding Member). Many at Aravind have noted that staying true to its mission is a crucial trigger that gets people up in the morning every day to perform: *"The organization is maintaining values through leaders playing the role-models. We become an example. For example I come here at 7 am in the morning, they come too even if there is no such rule"* (Chairman).

Functional Mechanism (5) – achieve operational productivity

A key set of functional mechanisms driving operational productivity at Aravind are a dedicated commitment to standardization, the provision of real time performance measures, and incremental experimentation. Everything required to deliver high quality eye surgery is standardized and coordinated. This involves the organization of eye camps for fast and efficient scanning of potential rural patients, optimization of logistics and transport of patients to the main hospital, the actual surgical procedures and the post surgical treatment (Rangan and Thulasiraj, 2007). Strict task specialization at every level of the organizational hierarchy enables steep learning curves and focused skill development. The pace of highly formalized processes triggers predetermined routine action. Time compression suppresses the ability to reflect on a situation and develop action alternatives which in an optimized system are likely to be dysfunctional: *“Every case is a replica of another case. The paramedical staff smoothes the work, the time lost between patients goes down. That is how the system runs. Paramedical staff, trained well, cut down the time. Performance level of the doctor is enhanced because his work is facilitated. A normal doctor would be able to do 6-8 surgeries an hour here. It’s all about time management”* (Chief Medical Officer).

However, task specialization resembling a Taylorist approach to human resource management can create dysfunctional mechanisms such as alienation, boredom, rebellious behavior and high turnover at all levels. This could prevent the regular production of desired outcomes. Aravind has various methods for disabling and suppressing such dysfunctional mechanisms. Nurses for example go through a uniformly structured two-year training program. Consequently, they have the same set of skills and can be replaced with minimal task interruption, which suppresses their incentives to rebel and disables the negative organizational impact of any nurse leaving. By providing uniform in-house training coupled with equal pay across task domains, Aravind is able to suppress sentiments of task inequality which could negatively affect commitment and work attitude and therefore compromise regularities in outcomes.

The extreme task specialization of surgeons might create fears of deskilling and loss of marketability. This could trigger the dysfunctional mechanism of losing highly trained eye doctors who are still bottleneck resources in India. Aravind’s structural response to suppress such a detrimental effect involves managing several channels for engagement with the international community of leading ophthalmologists. Furthermore, Aravind invests in facilities that enable cutting edge research in a number of eye disciplines. And the large number of surgeries at Aravind provides doctors with unique access to special or rare cases: *“... we also need to have this external interface. Otherwise, we get too inbred. [...] people take part in lots of conferences [...] Then we also have the residents from the best universities in the United States come as part of their rotation. All these processes, almost through osmosis, kind of brings in best practice from outside into the system”* (Executive Director).

A second enabling and triggering mechanism to enhance operational productivity is represented by the provision of stretch goals and the transparent and real time provision and exposure of task performance: *“Sometimes we even set some kind of targets, especially on quality, like pressure-reduced complication rates, or we should achieve outcomes of this standard. Then we have a fairly robust system for collecting information, analyzing, then reflecting on it. This is, I would say, a formal process, but which continually evolves as well, which you can reckon is more internally focused”* (Executive Director). In this case the process leading to outcome regularities spans across hierarchical levels, i.e., performance is reported and made public across task domains and hierarchies: *“We do a few things during the course of*

the year to kind of connect back to the founding values and principles of the organization. We have like competitions across the organization to kind of say how their work connects up with the mission of the organization, how this driving, or maintaining cars, or cleaning a toilet: how does it add up to reducing needless blindness” (Executive Director). As this quote illustrates, task performance is directly and explicitly related to the mission and organizational objectives, which allows Aravind to disable and suppress mission drift and trigger continuous commitment to task performance. Furthermore Aravind disables or suppresses possible dysfunctional mechanisms typically associated with competitive approaches to task performance. For example, during our repeated field visits we observed on several occasions that numbers of surgeries performed by each doctor are put on a wall at the end of each day. Although the name of the doctor is not openly revealed, everybody in the hospital is informed about the absolute and relative performance. Aravind explicitly avoids individualizing performance by for example, selecting a “surgeon of the month” or “nurse of the day”. This approach to reveal task performance suppresses a potentially dysfunctional winner/loser atmosphere without compromising the motivational effect of revealing performance levels: “Discipline is, I would not say it is a cultural value, but it is a process which is extremely important in a hospital setting, because it is a teamwork. For discipline, in terms of coming on time, or all the work. If those things are violated, then the whole system suffers. We kind of give a lot of weight to that” (Executive Director).

A third mechanism to enhance operational productivity is incremental experimentation to optimize all functional mechanisms. But experimentation is tightly controlled and monitored to avoid potential dysfunctional mechanisms, e.g. allocating scarce resources to doctors’ pet projects. Experimentation is further discussed in the section on scaling below.

Functional Mechanism (6) – achieve profitability

Profitability allows Aravind to be independent from fluctuating levels of donations that might be difficult to control with concomitant loss of closure. Furthermore, profitability is an essential means to stay true to Aravind’s mission of inclusive growth. Apart from high productivity, generating demand from paying patients is thus a key functional mechanism. Building capacity to accommodate paying patients, for example by establishing hospitals with state of the art design and technology as well as offering private accommodation options, enables demand generation. Yet, why would paying customers choose Aravind over any other private for-profit hospital? Aravind eliminates dysfunctional trends such as losing patients to competitors by offering slightly lower rates for paid surgeries, better food and superior post surgical service for paying patients. In addition to nurturing Aravind’s reputation for high quality surgery, the authenticity generated by sustaining an ability to treat 70% of its patients for free, has created a strong reputation for being a particularly caring organization that paying patients value: “Last year, we did about 5,000 camp surgeries. Among that, 2500 [were] totally free of cost. Among that 2500, about 1000 [were] self-sponsored (by our group). Sometimes, a paying patient would like to sponsor 4-5 patients. When they visit the hospital, we tell them that some part of their payment goes to cover free patients, they ask how much is the cost for a free patient, we say 700, they say ok, I want to pay for 5-10 free patients” (Director). Because of Aravind’s commitment to high volumes and high quality, its doctors are running down learning curves very fast and Aravind has a pool of the best eye surgeons of the world.

Implications for Scaling

We base the development of propositions on our analytical framework and on Aravind data. The propositions are contingent on a context of large-scale poverty that legitimizes the perspective of inclusive growth models. The way we conceptualize that context has implications for our perspective of scaling. In the context of poverty, resources of many kinds tend to be scarce and often are already employed to enable basic livelihoods. Any alternative uses of resources may therefore have high social costs. Innovation may not always be the most efficient and productive use of available resources in such a context (Seelos and Mair, 2007). But some relevant dimensions of the environment are stable and predictable. Significant demand is enabled, yet not necessarily triggered, by large numbers of people with fundamental needs that could be satisfied by simple products and services. Competition as a potential dysfunctional mechanism is mostly absent. The lack of competition over how best to serve customers explains the persistence of unfulfilled human needs on a mass scale. And it also creates a relatively predictable task environment for an organization willing and able to tap into the large-scale demand once the triggering mechanisms that transform needs into real demand are established. This eliminates competitive variance from the context and thus facilitates closure. Large-scale poverty therefore looks like an opportunity for organizations to be extremely efficient and to deliver highly standardized products and services on a mass scale for sustained periods and without much adaptation or the necessity for radical innovation. Because of low income levels, serving the poor on a large scale requires a strong focus on cost control and operational efficiency. This implies a particular need for organizational settings that enable focused experimentation and marginal but continuous improvement in addition to the need for innovation in the sense of creating new recipes.

Scaling as Increased Productivity

The perspective of organizational closure has a number of implications for scaling. As discussed, from a realist philosophy closure is a requirement for scaling where scaling means the increased achievement of concrete and ex-ante specified organizational outcomes. The link is specified by the available knowledge or recipe of the adequate set of functional mechanisms known to create organizational objectives as their effect. But that knowledge or recipe is not sufficient in itself because it only specifies what "could be" and not what "will be". The latter is contingent on a number of intervening or dysfunctional mechanisms that could prevent the results of operating functional mechanisms to be realized. Closure requires knowledge about the required set of functional mechanisms and the ways in which possible dysfunctional mechanisms can be disabled and suppressed. Following the analytical framework in Figure 1, we can conceptually distinguish "degrees of closure" where higher levels of closure correspond to better organizational abilities to enable and trigger functional mechanisms in such a manner that their effects are regularly realized. Higher degrees of closure can also be achieved by recognizing potential dysfunctional mechanisms and learning how best to disable and suppress them. The many ways in which Aravind achieves closure in the section above highlight this important link between closure and scaling. It also gives meaning and relevance to the concept of "closure competence" as an important resource for organizations striving for inclusive growth.

Proposition 1: Increased organizational closure is achieved by the generation of more functional mechanisms and less dysfunctional mechanisms and thus correlates positively with achievement of organizational objectives.

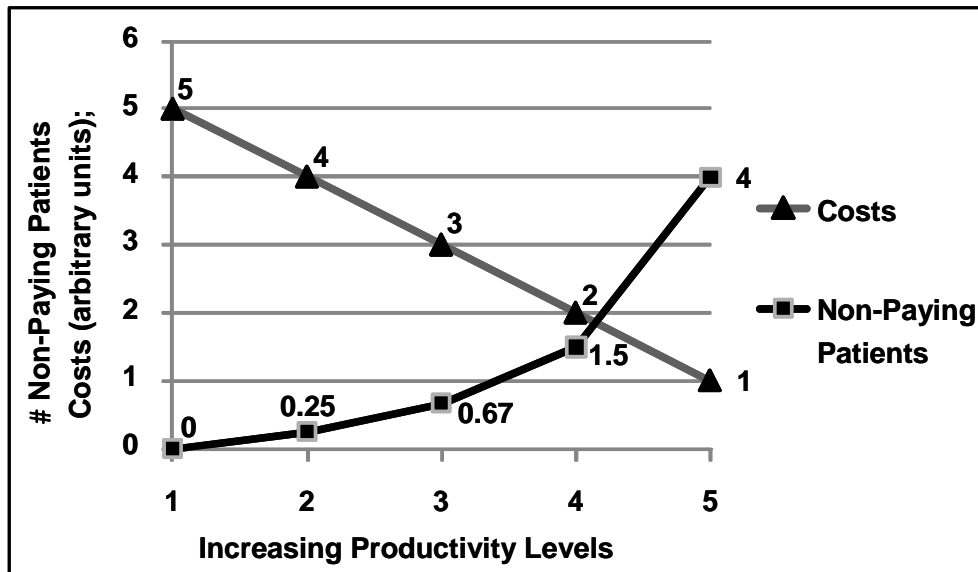
Knowledge about cause and effect mechanisms is a prerequisite for closure. Closure implies specification of the functional mechanisms required to achieve desired outcomes. We argue that the relation between knowledge and closure is an iterative process where better specifications of mechanisms enable higher degrees of closure. However, higher degrees of closure better resemble conditions that enable controlled experimentation. Controlled experimentation enables better isolation of the effects of newly introduced mechanisms and thus facilitates systematic learning. This lowers uncertainty about cause and effect relations, and strengthens the quality of the knowledge base or recipe. Performance thus improves by replacing functional mechanisms with substitutes that are more productive and by identifying and eliminating functionally neutral mechanisms or organizational slack. The latter is crucially important in environments of low munificence where many types of resources are scarce and their opportunity cost is high. Inclusive growth in such environments makes increasing productivity an essential scaling mechanism. *“The operating room staff meet every week. They would review the issues that they had. It could be sometimes some supply was not there when it was required, or suddenly some equipment failed, or there was a complication, or the patients sat way too long in the operating room. It could be any one of those things. Then, obviously, solutions to these problems are best with the people who are dealing with it. When they raise a solution and then you act on it, it makes things better. That is how we built an ownership for the process, and for the organization. When you do it on an ongoing basis, dealing with smaller and smaller problems, then you kind of move towards perfection”* (Executive Director).

Proposition 2: Closure increases productivity levels through a positive feedback mechanism between the quality of knowledge and higher degrees of closure.

Productivity increases have particular effects on performance in subsidized business models, which often are part of an inclusive growth strategy such as Aravind's. Given the constraint of financial self-sufficiency as an organizational objective, increasing the ability to serve poor "customers" that cannot pay market prices is determined by two factors: 1) increasing the numbers of paying customers to generate excess profits to subsidize poor customers, and/or 2) increasing productivity levels. A simple calculation demonstrates the effect. We use the Aravind model and assume that paying patients pay five units for a cataract surgery. Increased productivity is achieved by lowering marginal costs per surgery from an initial level of five units (with zero profits available to subsidize free surgeries) and decreasing cost levels to one unit per surgery which enables four free surgeries for every paid surgery at the financial constraint of zero loss. This relationship is shown in **Figure 3** below.

Figure 3

Relationship between productivity levels (defined as costs per treatment) and ability to fund treatments of non-paying patients per paying patient



For example, in the early 1990s Aravind decided to eliminate a dominant cost bottleneck – the highly priced intra ocular lenses required for cataract surgery. It founded Aurolab and thus brought the lens prices down from about 200 dollars to about 4 dollars. This greatly facilitated the ability to treat up to 50% of patients for free and still create the surplus required to fuel Aravind’s inclusive growth model. *“Everybody thought when you do free work you will not become viable financially and Dr V [Aravind’s founder] said, when you know your expenditure a little bit, when your productivity is better, when you have a good management system, you can be very viable. That’s our model. So the model became high volume, high quality, affordable cost. Volume was high so productivity became more and productivity became more and money came. When money came we could treat more people”* (Director).

Proposition 3: For profitable business models that also serve the poor despite lower marginal revenues, a marginal increase in productivity has positive scale effects on the numbers of poor customers that can be served.

Scaling as Increased Size

An obvious scaling mechanism is the addition of resources to an organization. Closure allows an organization to identify resources that enable functional mechanisms and the organizational design enables the triggering of these mechanisms and ensures the realization of outcomes by disabling dysfunctional mechanisms. Adding resources that enable functional mechanisms thus directly scales operations and achieves higher levels of outcomes. However, adding resources increases organizational size and this generates higher complexities that make it more difficult to maintain closure conditions. *“Lots of things are changing. As the system is getting bigger, you shouldn’t get diluted. Somehow, as things are managed here, it’s not diluted because of the close knit of the senior group that’s always there and that binds you together and that makes you do things in a certain way”* (Director).

Proposition 4: Adding resources that enable functional mechanisms to an organization that has achieved closure has a positive effect on performance up to the point where marginal loss of closure due to organizational complexity exceeds marginal gain in performance.

Scaling as Replication

Proposition 4 implies replication as a mechanism to overcome the liabilities of scaling organizational size. Replication allows the establishment of a copy of the original model. This creates an opportunity to replicate at an organizational scale that still has positive marginal performance when adding resources. However, in environments of low munificence, resource constraints of all kinds are a challenge to replication. Thus, lower productivity levels make it imperative to first create the knowledge that enables higher closure and therefore increased productivity, before replication can be expected to make a positive overall contribution to scaling.

Reflecting on the need to replicate the whole system of Aravind not just modules such as technology, a founding member commented: *“Aravind is very particular. [...] our success is technology, team work, commitment, hard work, willingness to change; and we have our own values and culture. So that also should calculate into the human resource area; and whatever we focus [on ... for example] the attitude of the person, everything should be similar otherwise there is no use of the best technology. Everybody can go and operate, everybody can do better surgery than us, have a better building than us. The model of high volume, high quality, affordable cost, with compassionate care to all, that is our own model, which we need to transfer.”* Furthermore, Proposition 2 implies that higher levels of closure facilitate the replacement of functional mechanisms by better substitutes. This means that replicated organizations have an opportunity to maintain closure across the branches and thus diffuse new mechanisms rapidly and in the same manner. However, the ability to achieve fast dissemination of new mechanisms also requires higher degrees of closure as there will be less local interpretation of the feasibility or necessity for change and the new mechanism is more likely to work as expected. Aravind has four branch hospitals which regularly exchange comprehensive reports on finance, surgery performance and quality. A director at Aravind reflected on this sharing mechanism: *“All the finance gets updated here in Madurai [...] almost every day. [...] For example, there was a remodeling done in the glaucoma clinic in Coimbatore [branch]. They reorganized the patient flow. They were able to bring down the unnecessary waiting time from 2 hours to 18 minutes. That was very dramatic. So, the design of the area and the process changed. And that spread like wood fire across the system. That sharing happens. It was shared in one of the clinical heads meetings.”*

Proposition 5: Given resource constraints, higher levels of productivity are positively associated with the organizational ability to achieve scale through replication.

Scaling as Knowledge Transfer

When organizations do not have the resources or do not want to replicate, knowledge transfer to other organizations is a potential mechanism for scaling. Our perspective implies that knowledge transfer cannot be expected to improve the performance of the receiving organization unless it has the closure competencies necessary to create desired outcomes as a result of integrating knowledge resources. An Aravind consultant reflected on the differences between managed care as a much more hands-on mode of knowledge transfer over longer periods of time and the work of Aravind consulting that focuses more on short-term engagements and exchange of best practices: *“In managed care, we can take some credit, when*

we are running it on a regular basis. But in consulting, we only act as a catalyst, we can't directly say 'it is because of me'. Because, once we act as a catalyst and they receive the ideas, they'll have to start implementing it from their side. But there are hospitals which have not done well because of lack of or frequent change of leadership or lack of availability of doctors or they're not following the processes they should" (Director). Even the much more hands-on managed care approach to knowledge transfer is challenged by the inability to replicate the organizational circumstances required for the knowledge or recipe to realize its expected outcomes: "In managed care, it's been a challenge of putting in the right structure, but we're able to put in the systems relatively well. But putting in the culture has been a challenge. I don't think we've succeeded in doing that" (Executive Director).

Proposition 6: Knowledge transfer from an organization that has achieved closure will not enable similar levels of performance in the receiving organization if it cannot establish the same degree of closure as the original organization.

Discussion and Conclusions

In many countries, poverty persists despite all efforts to replicate economic development. The need for innovation to develop new recipes able to integrate the poor and able to work in a given context is high. The objective of this paper was to pursue the question of how new recipes can be scaled. We see scaling as complementary to innovation. Both are required to achieve inclusive growth that integrates the poor and reaches service levels concomitant with the extent of poverty. The academic literature on scaling is in an embryonic stage. A recent review² lists fifteen key references. Of these only three are published in scholarly journals highlighting the interest and need of practitioners for a better understanding of the concept of scaling. We contribute to the literature by providing a scholarly treatment of scaling. First, we focus on an explicit definition of the concept. Scaling in this paper refers to an event regularity in the sense that doing more of A or doing A better creates more of an expected outcome, B. Second, we apply the perspective of critical realism postulating that event regularities require closure that is achieved only in artificial experimental situations. We therefore investigated the possibility of quasi-closure as an organizational characteristic required for generating sustained event regularities. Third, we identified an organization that has provided a highly standardized medical service over more than two decades as an extreme case of sustained event regularities. Reflection on original case data and theoretical literature enabled us to develop our central analytical framework. This framework constitutes a theoretical model of organizational closure. It facilitates a retroductive logic explaining how actors, structures, mechanisms and outcomes are coupled sustainably to achieve organizational objectives. Our perspective also highlights the need for a counterfactual explanatory logic. We retroduce the requirement for disabling or suppressing potential alternative mechanisms. This matters hypothetically insofar as their realization would constitute countervailing or dysfunctional mechanisms that would prevent the actualization of the effects of functional mechanisms and thus the achievement of objectives. Therefore, even if dysfunctional mechanisms are not actualized and thus would escape empirical investigation, an explanation of organizational closure requires consideration of how they are disabled or suppressed. We find that closure is an organizational competence required for our definition of scaling. Closure competence refers to a distinct set of organizational knowledge about: a) the required set of

² Scaling Social Impact Research Project (2006) CASE, Fuqua School of Business, Duke University.

functional mechanisms to achieve objectives; b) how actors and structures need to be selected and configured so that functional mechanisms are enabled and triggered on a sustained basis; c) the expected set of dysfunctional mechanisms that actors and structures could generate in parallel or as an alternative to functional mechanisms, and d) how to disable or suppress dysfunctional mechanisms so that the effects of functional mechanisms can be repeatedly and reliably actualized. Finally, we developed propositions about scaling that followed from our analytical framework and reflections on the illustrative case study.

Implications for pPractitioners

Our propositions on scaling mechanisms highlight the importance of closure to achieve productivity. Given the scarcity of resources in the context of poverty, productivity measures, i.e. outcomes generated from a given resource, are essential metrics for project evaluation. Funders of inclusive growth models typically evaluate projects by isolated input or output measures. The perspective of productivity requires an explicit evaluation of the coupling between input and output variables. Our framework on closure provides practitioners with a systematic view on productivity. It also facilitates a more differentiated perspective on the types of resources required for inclusive growth. Often, monetary resources, e.g. development funds, grants or donations are considered crucial factors for progress. However, our framework helps to identify critical resources based on identification of the functional mechanisms they are able to generate. Money usually cannot buy these resources in the context of poverty and thus long-term efforts to provide bottleneck resources need to receive explicit attention. Aravind made a key investment in its own training program for productive eye doctors, a real bottleneck resource given its inclusive growth model. This has implications for rethinking funding strategies and to make more realistic evaluations of development projects (Pawson and Tilley, 1997).

Our scaling propositions allow us to illuminate the link between resource availability and organizational closure. This link gives rise to several scaling options that may support decision making for managers as well as funders. When resource levels and degrees of organizational closure are low, the best option is to enhance productivity by learning how to increase closure. Where both degrees of closure and resource levels are high, understanding key resources and adding capacity is the prime scaling option. If organizational size reaches a point where complexity causes a loss of closure the best option is to invest in replication of the model at another location. Our analytical framework supports replication decisions by enabling a sophisticated understanding of contingency issues, i.e. to what extent actors and structures in a new location enable functional mechanisms and/or may undermine project success by generating too many dysfunctional mechanisms. Where degrees of closure are high but levels of resources are low knowledge transfer to another organization particularly in a more resource abundant location is a valid scaling option. In addition, investments in building key bottleneck resources are necessary to improve capacity of the original organization. When degrees of closure are low, as when organizations have not found solid working recipes and when levels of resources are high, the best option is to invest in further innovation and prototyping. Higher levels of resources are compatible with the uncertainty and failures implicit in innovation. Emphasis should be on critical evaluation of what works and what does not in order to limit allocation of resources to inefficient and ineffective recipes. **Figure 4** summarizes these implications.

Figure 4

Scaling options contingent on levels of resources and degrees of organizational closure

Degrees of Closure		
Low	High	
<ul style="list-style-type: none"> •Experimentation to increase productivity 	<ul style="list-style-type: none"> •Knowledge Transfer •Development of key resources 	Low
<ul style="list-style-type: none"> •Innovation and prototyping to create better recipes 	<ul style="list-style-type: none"> •Adding resources to increase capacity •Replication 	High

Implications for Research

Despite its promise, critical realism has received limited explicit attention in research on strategy and organizations. Tsoukas (1989) has argued for the epistemological validity of idiographic case research from a critical realist tradition. With this study we hope to demonstrate the usefulness of this perspective for investigating the concept of scaling and to inspire others to expand and sharpen our analytical framework and the ways it can be used for theorizing.

We provide a contrasting perspective on scaling based on a narrow definition grounded in the perspective of closure. Others look at scaling from an open systems perspective that includes multiple interactions with a variety of stakeholders which expands the scale and scope of activities. It also includes evaluating outcomes much more broadly by focusing on indirect impacts such as spill-over effects as well as unintended consequences and different dimensions of value creation (democracy, empowerment etc). This is perhaps best reflected in a quote by Uvin (2000, p. 1418): *"In the new paradigm, the extent to which an NGO successfully scales up can be judged not only in terms of its size, but also in terms of the number of spin-offs it created, the number of projects that have been taken over by other actors, and the degree to which it contributed to the social and intellectual diversity of civil society. [...] Impact, finally, is not only about the number of beneficiaries or even the specific policy changes won, but also about local capacity built, intersectoral contacts developed, norms of trust and cooperation strengthened, and democratic space and social diversity reinforced."* Therefore, our definition of scaling limits the generality of our findings to one particular meaning of scaling and to a unique context of large-scale poverty. Much more systematic and comparative research is required to improve our understanding of the contingent nature of the interplay between functional and dysfunctional mechanisms in their ability to create empirical outcomes.

Finally, we wish to share our experience applying critical realist perspectives as stimulating very rich discussions amongst the research team and encouraging deep thinking which we found quite challenging and enjoyable. It made us reflect deeply on the usefulness and limitations of mainstream research methods and our own commitments to ontology and epistemology – an ongoing inquiry into our roles as researchers.

References

- Ackroyd, S. and Fleetwood, S. (2000), "Realist perspectives on management and organizations," New York: Routledge.
- Archer, M. (1995), "Realist Social Theory: The Morphogenetic Approach," Cambridge: Cambridge University Press.
- Archer, M. S. (2002), "Realism and the problem of agency," *Journal of Critical Realism*, 5, pp. 11-20.
- Baum, J. A. C. and Rowley, T. J. (2005), "Companion to organizations: An introduction," in Baum, J. (ed.), "*The Blackwell Companion to Organizations*," Oxford: Blackwell.
- Bhaskar, R. (1975), "A Realist Theory of Science," Leeds: Leeds Books.
- Bradach, J. (2003), "Going to Scale: The Challenge of Replicating Social Programs," *Stanford Social Innovation Review*, 1, pp. 18-25.
- Bunge, M. (2006), "Chasing Reality: Strife over Realism," Toronto: University of Toronto Press.
- Chen, S. and Ravallion, M. (2007), "Absolute poverty measures for the developing world, 1981-2004," *Proceedings of the National Academy of Sciences*, 104 (43), pp. 16757-16762.
- Collier, A. (1994), "Critical Realism," London: Verso.
- Commission on Growth and Development (2008), "The Growth Report. Strategies for sustained growth and inclusive development," Washington: The World Bank.
- Davies, J. B., Sandström, S., Shorrocks, A., and Wolff, E. N. (2008), "The world distribution of household wealth," *UNU-WIDER Discussion Paper 2008/03*.
- De Véricourt, F. and Sousa Lobo, M. (2009), "Resource and revenue management in nonprofit operations," *Operations Research*, 57, pp. 1114-1128.
- Durand, R. and Vaara, E. (2010), "Causation, counterfactuals and competitive advantage," *Strategic Management Journal*, 30, pp. 1245-1264.
- Easterly, W. (2002), "The Elusive Quest for Growth," Cambridge, MA: MIT Press.
- Easton, G. (2000), "Case research as a method for industrial networks. A realist apologia," in Ackroyd, S. and Fleetwood, S. (eds.), "Realist perspectives on management and organizations," pp. 205-219, New York: Routledge.
- Fairclough, N. (2005), "Discourse analysis in organization studies: the case for critical realism," *Organization Studies*, 26, pp. 915-939.
- Harré, R. and Madden, E. H. (1975), "Causal Powers: A theory of natural necessity," Oxford: Blackwell.
- Hodgson, G. M. (2007), "Institutions and individuals: Interaction and evolution," *Organization Studies*, 28, pp. 95-116.

- Johnson, P. and Duberley, J. (2000), "Understanding Management Research: An Introduction to Epistemology," London: Sage.
- Lawson, T. (1995), "A Realist perspective on contemporary 'Economic Theory,'" *Journal of Economic Issues*, 29, pp. 1-32.
- Lawson, T. (1997), "Economics and Reality," London: Routledge.
- Leca, B. and Naccache, P. (2006), A critical realist approach to institutional entrepreneurship," *Organization*, 13, pp. 627-651.
- Pawson, R. and Tilley, N. (1997), "Realistic Evaluation," London: Sage.
- Rangan, V. K. and Thulasiraj, R. D. (2007), "Making sight affordable," *Innovations*, 2, pp. 35-49.
- Reed, M. (2005), "Reflections on the realist turn in organisation and management studies," *Journal of Management Studies*, 42, pp. 1621-1644.
- Rodrik, D. (2006), "Goodbye Washington Consensus, Hello Washington Confusion? A review of the World Bank's economic growth in the 1990s: Learning from a decade of reform," *Journal of Economic Literature*, 44, pp. 973-987.
- Sayer, A. (1992), "Methods in Social Science. A Realist Approach," 2nd ed. Routledge, New York.
- Seelos, C. and Mair, J. (2005), "Social entrepreneurship: Creating new business models to serve the poor," *Business Horizons*, 48, pp. 241-246.
- Seelos, C. and Mair, J. (2007), "Profitable business models and market creation in the context of deep poverty: A strategic view," *Academy of Management Perspectives*, 21, pp. 49-63.
- Thompson, J. (1967), "Organizations in Action," New York: McGraw Hill.
- Tsang, E. W. K. and Kwan, K. (1999), "Replication and theory development in organizational science: a critical realist perspective," *Academy of Management Review*, 24, pp. 759-780.
- Tsoukas, H. (1989), "The validity of idiographic research explanations," *Academy of Management Review*, 14, pp. 551-561.
- Tsoukas, H. (2000), "What is management? An outline of a metatheory," in Ackroyd, S. and Fleetwood, S. (eds.), "*Realist Perspectives on Management and Organizations*," pp. 26-44, New York: Routledge.
- Uvin, P. (1995), "Fighting hunger at the grassroots: paths to scaling up," *World Development*, 23, pp. 927-939.
- Van de Ven A. (2007), "Engaged Scholarship: A Guide for Organizational and Social Research," New York: Oxford University Press.
- Whitley, R. (1984), "The scientific status of management research as a practically-oriented social science," *Journal of Management Studies*, 21, pp. 369-390.
- World Health Organization. (1997), "WHO Fact Sheet N143," Geneva: World Health Organization.
- World Health Organization. (1999), "WHO Fact Sheet N230," Geneva: World Health Organization.

Appendix A

Description of the Interview Data Collected

Number of hospitals visited	6 (3 branches: Madurai, Pondicherry, Coimbatore; 3 affiliated units: Lucknow, Amethi, Kolkatta)
Number of interviews conducted	51 (14 of the interviews involved teams or groups, e.g., the management team and groups of nurses)
Number of interviewees involved	63
Organizational roles of the informants	<ul style="list-style-type: none">• Chairman of the organization• Founding members• Executive director• Directors and managers of sub-units• Finance, infrastructure, training managers• Hospital administrators• Medical officers• Doctors• Nurses• Trustee representatives, external collaborators, members of supporting institutions
Duration of the interviews	10 min – 1 hour 40 min
Number of pages of interview data	More than 350 pages
