INDICATORS AS INTERVENTIONS:

PITFALLS AND PROSPECTS IN SUPPORTING DEVELOPMENT INITIATIVES

A report prepared with support from the Rockefeller Foundation

KEVIN E. DAVIS and BENEDICT KINGSBURY

December 2011

About the authors

Kevin Davis is Beller Family Professor of Business Law at New York University School of Law.

Benedict Kingsbury is Murry and Ida Becker Professor of Law, Director of the Institute for International Law and Justice at New York University School of Law (iilj.org), and Visiting Professor of Law at the University of Utah.

Acknowledgements

This report draws extensively on background papers authored for this project by Nehal Bhuta, Sarah Dadush, and Angelina Fisher. Their papers appear, with substantial modifications, in K. Davis, A. Fisher, B. Kingsbury, and S. Merry eds., *Governance by Indicators: Global Power Through Quantification and Rankings* (Oxford University Press, forthcoming 2012). The authors of this report have also benefited greatly from an incisive review of relevant literature prepared by Christopher Bradley, and from drafting suggestions and comments provided by Angelina Fisher and Chris Bradley and by a group of interlocutors at the Rockefeller Foundation.

The material in this report also draws heavily on research the authors have conducted jointly with Sally Merry. The key ideas are presented in K. Davis, B. Kingsbury and S. Merry, "Indicators as a Technology of Global Governance," 46 *Law & Society Review* (2012) and in the book *Governance by Indicators*. Davis, Kingsbury, and Merry are currently leading a multi-country collaborative project focused on indicators in developing and transitional countries, funded by the National Science Foundation. These other research projects have been generously supported by Carnegie Corporation of New York, the National Science Foundation, the Straus Institute for the Advanced Study of Law and Justice at NYU, and NYU Law School's D'Agostino and Greenberg Faculty Research Fund.

TABLE OF CONTENTS

Foreword i

Executive Summary ii

1. Introduction 1

1.1. Top-down problem-solving and the pathologies and pitfalls of indicators 1

1.2. Indicators as interventions aimed at wicked problems 2

1.3. Overview of the report 4

2. Indicators and their alternatives 6

2.1. What is an indicator? 6

2.2. How indicators are produced 8

2.3. Alternatives to indicators 10

2.4. Case Studies 10

1. WHO/UNICEF Immunization Coverage Indicators 11

2. State Failure: The U.S. Fund for Peace Failed States Index 12

3. Impact Investment: The Global Impact Investing Rating System (GIIRS) 13

3. Roles for indicators in addressing social problems 15

3.1. The process of addressing social problems: framing, action, contestation, updating 15

1. Framing 15

2. Action 16

3. Contestation 16

4. Learning and Revision 16

3.2. Two roles for indicators 16

4. Indicators and the framing of social problems 19

4.1. How indicators frame problems 19

4.2. Indicators: validity and measurement error 20

4.3. Discrepancies in framing between the indicator and the gold standard for the problem 21

4.4. Indicators that frame problems controversially 22

4.5. Do indicators promote common understandings? 24

5. How indicators influence action 25

5.1. Which indicators influence action? 25
5.2. What factors determine the influence of indicators? 27
5.3. Under what conditions can indicators promote optimally constructive action? 29
5.4. When do indicators promote coordinated action? 31
5.5. Does use of indicators enhance accountability for actions? 31
6. Contestation around indicators 33
7. Learning and revision 35
8. Conclusion 37
Appendix 39

Foreword

Improving the well-being of poor and vulnerable people relies on solving systemic problems. The perceived severity of these problems and the success of potential solutions are often measured by information bundled into key indicators. Gaining a clearer understanding of the role that indicators play is a critical issue that will only take on more importance as new information become increasingly available and used over the next decades. To help catalyze innovation and improve our collective ability to have impact, the Foundation has provided support for this case study research on indicators to inform thinking and practice in the development sector and beyond. This is an important topic for the Rockefeller Foundation, which has been involved in creating and disseminating a number of indices throughout its history. For example, the Global Impact Investing Rating System (GIIRS), supported by the Foundation and covered as a case study in this report, is a standard-setting component central to building an infrastructure for the impact investing industry that can facilitate the scale-up of the marketplace for institutional and individual investors seeking to address social or environmental problems while also maintaining capital growth. Throughout our partnership with the non-profit organization B Lab, we sought to reframe indicators as an intervention rather than just information and ask whether new insights could emerge that provide guidance for practitioners.

One of the most interesting findings presented in the report is that the process of creating and disseminating indicators can be an effective intervention that is particularly useful in addressing and marshaling a response to wicked problems—complex, interdependent, ever-changing global issues that require the application of iterative solutions in order to be managed successfully. As the authors point out, it is valuable to compare indicators with other potential interventions in the international development system and, in their most compelling and powerful form, use them to trigger actions that move us one step closer to addressing challenges affecting the lives of poor and vulnerable populations on a daily basis.

The ability of indicators to help reframe problems related to poverty and globalization is an excellent example of how new kinds of evidence can play an outsized role in shaping the responses undertaken by philanthropies, governments, and other organizations interested in ensuring social change. Moving forward, the growing emergence of user-generated information that actively involves beneficiaries in the collection, production, and assessment of data will likely significantly shape the next generation of indicators. Similarly, a shift in the locus of indicator construction is also likely to take place in the coming years, moving away from the current situation in which institutions based in the global North produce indicators about challenges taking place in the global South and, instead, lead to a rise in South-South collaboration related to indicator construction. An example of this trend is the Ibrahim Index of African Governance created by the Mo Ibrahim Foundation, an instance where an African-based institution has developed an increasingly well-known indicator that tracks the quality of governance across the continent.

In conclusion, this important contribution to the field highlights how indicators create valuable, necessary, and quantifiable simplifications that both illuminate key dimensions of a complex problem while simultaneously allowing important comparisons to be made. Readers will find the report useful whether they plan to create a new indicator, want to better understand tradeoffs between indicators and other intervention options, or evaluate in what ways indicators can be deployed most effectively. Finally, the report highlights pathways for needed future research about how indicators can lead to action and impact.

Claudia Juech
The Rockefeller Foundation

Evan Michelson
The Rockefeller Foundation

Executive Summary

Introduction

Indicators are ubiquitous in development policy. Corruption. Human trafficking. Obstacles to doing business. Child mortality and morbidity. There are indicators for all of these social problems. These indicators are not only methods of selecting, evaluating or publicizing development interventions – they are interventions in their own right!

Indicators have potential to be constructive interventions in addressing *wicked* or even *super-wicked* problems of development. Wicked problems have several characteristics which make them challenging, but which the use of indicators may help (somewhat) in addressing. Wicked problems are difficult to understand in single standard categories – for example, state failure is variously addressed by external agencies as a development problem or a security problem, and local people may see it in other ways. Wicked problems are complex: they typically have many causes, and potential responses often involve a myriad of different actors. Moreover, these problems take different forms in different places, making each version of the problem somewhat unique. The diverse actors whose efforts are needed to fruitfully address a wicked problem do not see the problem the same way or with the same parameters; most may identify with only some aspects; each may pursue wider ideals or agendas that unwittingly pull away from constructive solutions; they may hold collectively incompatible views due to responding to different questions; they may not see themselves as engaged in any way in the same problem as some of the other relevant actors. Well-crafted indicators, carefully deployed, may help to overcome some of these challenges. Conversely, indicators that are poorly crafted or used with insufficient consideration may exacerbate the difficulties.

Drawing on three case studies and a review of the literature, this report provides a framework for understanding the effects of producing indicators in global governance and transnational contexts relevant to development, drawing attention to some pitfalls and perverse effects that can easily result from producing and using indicators without realizing that they are themselves interventions, or without sufficiently far-reaching thought and consultations.

What is an indicator?

In technical terms, an indicator is *a collection of named, rank-ordered, simplified and processed data that purports to represent the past or projected performance of different units*. An indicator simplifies and processes data about a named social phenomenon in a way that makes it possible to compare and evaluate units such as countries, communities, organizations, or individuals.

Indicators can be used to monitor and evaluate the extent of problems and the success of interventions. The can also have a community mobilizing effect, bringing attention to a

problem, drawing in actors from disparate areas of interest and expertise, and allowing these actors to develop shared understanding of the problem, share information about past or anticipated actions and cooperate across a range of activities. Indicators can also provoke institutional action, including national policies and government action -- or guide prioritization and allocation of funds. Indicators may also inspire critique, resistance, and demands for reform.

Examples of Indicators
The Political Risk Index, an indicator based on 12 political and social variables that is produced by Political Risk Services, a for-profit entity, is designed to be used by investors in deciding where to invest and what risk premiums to charge. In addition, one or more Political Risk Services sub-indicators form a component of each of the World Bank Institute's governance indicators. (*Source:* www.prsgroup.com/icrg.aspx)
The Africa Media Barometer, measures press freedom through assessments by media practitioners and civil society in each country using 45 indicators. The media project of the Friedrich-Ebert-Stiftung (FES) in Africa orchestrates this work and uses the information to raise public awareness for media protocols and declarations signed by governments but not implemented through national law, raise awareness among NGOs of the importance of media issues for "good governance", and provide the basis for a reform agenda to media NGOs. (*Source*: http://fesmedia.org/african-media-barometer-amb/)
Community Indicators Victoria aim "to support the development and use of local community wellbeing indicators in Victoria, Australia, with the purpose of improving citizen engagement, community planning and policy making" by providing a community wellbeing indicator framework, measuring everything from crime to household waste recycling to education levels to work-life balance. The indicators are intended to inform local policy makers' choices of initiatives, but they are also a means for social organizing and community education. (*Source*: http://www.communityindicators.net.au)
The Country Policy and Institutional Assessment (CPIA), an indicator of the quality of countries' policies and institutions promulgated by the International Development Association (IDA) – the branch of the World Bank responsible for providing financing on a concessional basis – is an important factor in determining the IDA's allocation of aid across eligible countries (other factors include population, Gross National Income per capita and the recent performance of World Bank projects in the country). The CPIA assesses countries on the basis of 16 criteria, which fall into four subcategories: economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions. The final category, public sector management, is also known as "governance." For each of the 16 criteria, a country is given a score ranging from 1 (low, or weak) to 6 (high, or strong). (*Source*: www.worldbank.org, CPIA questionnaire 2010)

Production of an indicator is often a collective process. Promulgators may attach their names to indicators to which a number of actors have contributed, lending a brand name to the collective product of a global information supply chain. The data collection process can also involve a large network of actors, while the methods used to process the raw data typically rely on contributions from some segment of the scientific community.

Steps in production of indicators:

- collecting and processing data
- giving a name and imprimatur to the indicator (promulgation)
- distributing it or making it accessible to the public, and publicizing its existence.

Production is a collective process with a diverse set of participants

In thinking about interventions that can effect social change, indicators should be compared to alternatives, including to interventions that go beyond merely providing data (e.g., provision of money, goods or services; capacity-building; regulatory change; litigation; etc.). It is also important to consider the interactions between indicators and other interventions: do they complement each other? Compete with each other? Act as substitutes?

Examples of Alternatives to Indicators:

- Non-numerical data such as photographs, videos, songs, and slogans
- Combinations of non-numerical data and numerical data: for example, the U.S. State Department's reports on human trafficking include numerical country rankings, and a set of vignettes about particular individual victims expressed in narrative form
- Legal processes and verdicts
- Qualitative assessments and reports
- Systematized searchable raw data

Case studies

The report draws on three in-depth studies of specific indicators:

- Immunization coverage indicators produced by the World Health Organization and UNICEF, which measure the number of children vaccinated as a proportion of the target population (Angelina Fisher)
- The Global Impact Investment Rating System, GIIRS being developed by B Lab, a non-profit based in Pennsylvania to "assess the social and environmental impact (but not the financial performance) of companies and funds using a ratings approach analogous to Morningstar investment rankings or S&P credit risk ratings" (Sarah Dadush)

- The Failed State Index, an index of state failure produced annually by the US Fund for Peace (USFfP), and published in *Foreign Policy* (Nehal Bhuta).

In addition to focusing on the specific indicator, each of the three cases also examined the roles that other indicators, as well as alternatives to indicators, play in addressing the problem that defines the case.

Role of indicators in addressing social problems

The process of addressing social problems can be modeled in four stages: *framing* of the problem, *action(s)* in response, *contestation* of approaches, and *learning and revision* resulting in rethinking or updating of the initial problem framing and selected actions.

Indicators can play a role at each of these stages. They do this both by serving as sources of information about a problem and by fostering social and political engagement.

In their roles as sources of information indicators can change some of the beliefs on which actors draw to address social problems, thereby affecting the decision-making process, actions, and eventually outcomes. For example, a state failure indicator that ranked Chad close to well-known failed states such as Somalia had "shock value" for readers unfamiliar with the deteriorating political situation in Chad. These kinds of shocks can be used to stimulate action.

In their socio-political roles, indicators may spark the creation of networks or organizations devoted to producing, responding to, or criticizing indicators. Those networks or organizations may in turn allow previously-unconnected people to develop new and shared understandings of how to frame problems, learn about causal relationships between actions and outcomes, share information, and cooperate.

Indicators and the framing of social problems

Indicators inescapably frame problems – they make statements about the existence and nature of a problem, as well as about *how to measure* the problem or aspects of its solution. For example, social impact indicators identify the social impact of business as a social concern, and embody claims about what count as more or less positive social impacts as well as how to measure them.

The name of an indicator may be determinative of the problem it will be used to frame. An indicator of state failure embodies claims about whether there is such a thing as a failed state, what it means to say that a state is more or less of a failure, and how to measure the extent of state failure. However, an indicator can serve to frame other problems besides the ones suggested in its name. For example, "immunization coverage indicators" have been used by funders as proxies for the overall quality of countries' health care systems.

By identifying factors that contribute to a problem an indicator can also implicitly embrace a specific *theory of change* – a theory about (i) what *causes* the problem, (ii) how those causal factors can be *overcome*, (iii) who the *key actors* are for the problem to be addressed productively, (iv) what *pathways* are likely to produce better outcomes, and (v) what would constitute *progress* toward solving the problem.

Indicators will succeed in influencing how people frame problems if people come to accept the indicator as the "standard" or "appropriate" way to measure the extent of the problem. Indicators also influence how people understand problems by providing starting points for dialogue. The process of deliberation and contestation can begin during the design and production of the indicator. Disparate actors may initially frame the problem very differently, thus triggering a process of deliberation over how to construct the indicator. This process is likely to change the views of at least some of the participants. Similarly, disseminating the indicator to a broader audience often prompts objections to the framing embodied in the indicator, with the ensuing debate leading at least some of the participants to reconsider how they frame the problem.

In the international context the use of standardized indicators to define and frame problems may result in the reshaping (or overtaking) of local understandings. This process is lauded or decried from different standpoints. However, there is no guarantee that the production of an indicator will in fact lead to common understandings of social problems. If each member of a group of users treats an indicator as just one of several sources of information about how to frame a problem, the group may end up with quite divergent understandings of the problem.

Some indicators frame problems better than others in that they provide more accurate information than others about the problems they purport to measure; more accurate indicators ordinarily provide better framings of the problem than less accurate ones. Indicators can lack validity, meaning that they attach misleading labels to data. Whether an indicator is valid depends upon the theoretical and empirical connections between the indicator and the phenomenon it claims to measure. The validity of an indicator can change over time, especially if the actors whose performance is being measured change their behavior in efforts to game the indicator. For example, a social impact indicator that is valid when initially designed may become less valid as firms learn how to boost their social impact ratings rather than actual social impact.

Inaccuracy can also stem from *measurement error*. These errors can be significant. Measurement errors can stem from a variety of factors, including *unreliable or missing data*. For example, a number of studies have challenged WHO "immunization coverage indicators" on the grounds that nationally aggregated data conceals geographic, gender, and socio-economic inequities in vaccine administration, and also hides localized sources of data irregularities.

Discrepancies in framing between the indicator and the gold standard for the problem

The possible relations between the way a particular indicator frames a problem, and the way a problem ought to be framed from the subjective standpoint of one observer or from the more objective standpoint of an existing 'gold standard', can be mapped two-dimensionally. Five different relations are possible, in this view.
The indicator may, in relation to the other (subjective or gold-standard) framing, frame the problem:

- Identically
- More narrowly
- More broadly (but fully subsuming the problem)
- In an overlapping way (subsuming part of the problem and also other separate issues)
- In an irrelevant way (addressing only separate issues).

Indicators are simplifications. It is thus extremely difficult to produce indicators that frame complex wicked problems in a way that is *identical* to an independent 'gold-standard' framing of the problem, especially if the gold-standard framing is not quantitative.

Indicators that are framed too *narrowly* can begin to define what they were intended only to proxy. The Human Development Index, for example, becomes equated with "human development" just as the GDP/capita becomes "economic development." Nonetheless, successful indicators typically embody narrow framings. Several indicators may be used, each of which deploys a different narrow framing (and these different narrow framings may overlap). For example, an immunization coverage indicator based on the number of children vaccinated frames the problem of under-immunization in a different way from an indicator that measures incidence of vaccine-preventable diseases among children.

Indicator may frame the problem more *broadly* than needed, in a way that captures several narrower framings. This can occur particular with composite (or mash-up) indicators. Take the case of indicators of state failure. A failed state can be understood as either a state that cannot promote the development of its population or as one that generates a risk of local, regional or global security threats. Interestingly, some failed state indicators provide data on both economic development and risks posed to other states, thereby framing the problem of state failure so broadly as to conflate the two narrower understandings. Lack of precision may undermine the indicator's value as an accurate representation of a problem and may limit its value in tracing the effects of interventions on the extent of a problem. For example, there is a real question of the extent to which the "failed state index" is useful (if at all) in addressing either institutional underdevelopment or security threat concerns underlying "failed states" because its framing of the problem is overly broad.

Indicators may also address part of a problem but at the same time *incorporate non-relevant material*. Mis-framings can cause or reinforce a mis-diagnosis of the problem and its causes, and mis-design of optimal interventions. In practice, however, it is often impossible to avoid some mis-framing, especially where indicators rely on data compiled for other purposes.

Some indicators purport to frame a problem, but in fact do not really address at all the problem they purport to address. This can readily occur where the data used is only

spuriously relevant to the real problem.

How indicators influence action

Indicators stimulate and shape action by alerting people to the existence of a problem, helping them to understand its magnitude, and pointing toward appropriate interventions. In addition, becoming involved in the process of creating or contesting an indicator can mobilize actors to focus on and address a problem.

Among our case studies, the indicators that most clearly exert this kind of influence are the WHO/UNICEF immunization coverage indicators, which are used by donors to decide how to allocate funding. The Global Alliance for Vaccines and Immunization (GAVI), for example, uses DTP3 coverage threshold to determine the type of funding available to a country. Under GAVI's rules, countries with less than 70% DTP3 coverage are *not* eligible for funding to support introduction of new vaccines. The privately-produced USFfP state failure indicator does not appear to have had a systematic impact of that kind, but preliminary evidence points to some use of the indicator (as one element) in decisions and actions. The social impact investment indicators we studied were too new to have had much impact on investment markets. But the process of producing them created a network of businesses and NGOs that may serve to expand interest in social impact investment, and their availability and further elaboration is likely to enable market activity by investors and aspiring recipients of investment.

Our work on the case studies reinforces a key point from the literature review conducted for this project: despite the vast effort that goes into production of indicators and growing frequency of their use, very little is known about the influence of most development-related indicators on decisions and actions. As the case studies demonstrate, it can be difficult even to determine the extent to which data from any given indicator is received by particular individuals; and if that is overcome, determining the extent of its influence on their ultimate actions is much more difficult. Actors often rely on many different kinds of data in the course of deciding whether and how to address problems, and they do not necessarily understand, record or disclose the elements of their decision making processes. Although the methodological problems for determining the impact of an indicator can be challenging, a conclusion of this study is that careful work to build evidence on these questions has become essential.

Conventional economic theory suggests that demand for any given indicator as a guide to action will depend on factors such as its accessibility and accuracy, the availability of information that substitutes for the indicator (assuming it is accurate, an indicator that provides different or "surprising" information that challenges pre-existing knowledge should be in relatively greater demand), availability of complementary products such as customized analytical tools, and whether use of the indicator generates network or learning effects (that is, their use by one user also provides benefits or insights for other users). For example, the DTP3 "immunization coverage indicator" became widely used as a proxy for health care performance because it is available for very many countries and sub-national areas with a

sufficient degree of standardization of methods and measures, cheap to access, has enough determinants to make it a reasonable proxy for health system performance, and there is no alternative single global indicator that measures national health system performance.

Indicators have the potential to point toward constructive action. Accurate indicators help actors identify the most constructive ways of addressing problems. The reverse is also true, inaccurate indicators point toward less constructive actions. Accurate social impact indicators will encourage socially-motivated investors to allocate funds to firms whose operations have the greatest social impact. Inaccurate indicators, however, will channel funds toward firms that have less positive impact than firms that are neglected.

Whether indicators have a constructive influence depends in part on the motivations of producers and users. To the extent that indicators provide information about the future, they can be used to profit from rather than to change the course of events. Investors might use an indicator of state fragility to decide which countries to avoid rather than to identify regimes that might benefit from constructive engagement.

Indicators have the potential to make decision-making more consistent and transparent. Relying on indicators will foster consistency in decision-making to the extent that indicators foster common understandings of a problem. But consistency in decision-making is not an unalloyed good. If a large number of major social impact investors use the same indicator to determine which firms are eligible to receive funds the result in an imperfect market may be a stampede to provide capital to a handful of firms, flooding highly rated firms with more capital than they can use.

Indicators also may contribute to the transparency of decision-making process because they are easier to communicate than other less heavily filtered data. Ease of communication implies that it is easier for people to access the direct informational basis for decisions—it is easier to communicate a country's score on a failed states indicator than it is to transmit the in-depth multi-faceted analysis of its political climate on which the indicator is based. The transparency of decision-making based on indicators, however, may be merely superficial if the process of producing the indicator is opaque.

Contestation around indicators

Indicators that are interventions are potentially likely to attract contestation. Broadly, if the indicator seems likely to have significance for what is valued or for future policies and resource allocation, people who do not share the framing, the causal theories, or the set of desired actions embodied in the indicator may protest it or create competing indicators and alternative framings.

The ability of different parties to effectively contest indicators will depend on who is best placed to produce and contest knowledge based on indicators according to financial resources and social privilege. Actors with superior abilities to collect, process, distribute and publicize the existence of data will generally be in the best position to produce

indicators. These are typically organizations with financial resources or social connections that give them privileged access to certain sources of data, ready access to social scientific expertise, and high public profiles. It is less clear however, which actors are best placed to contest knowledge based on indicators. Often a new indicator will be developed to challenge an existing indicator, making the ability to produce influential indicators an important characteristic that could shrink the group of contestants to those who are capable of producing indicators effectively.

Learning and revision

The data incorporated into an indicator itself may provide one basis for learning. The use of the indicator to frame problems and guide action, as well as any contestation it might provoke, all provide experience and ideas that form a further basis for learning. This kind of learning can propel revisions of framings of issues, plans of action, and practices and objectives of contestation. Moreover, inter-unit comparisons enable one unit to learn from others whose performance scores higher. Countries may learn from others about successful immunization strategies. Companies may learn from other more highly –rated companies in the same sector about techniques that work to increase ratings or attract investors. By contrast, it does not seem likely that states facing or experiencing failure learn much from indicators showing other states doing better or worse, although early warning may be valuable.

In principle, indicators themselves can be revised over time to reflect new framings, new ideas about causes of problems and appropriate interventions, and learning from contestation of the indicators or of theories embodied within them. The case studies in this report do not involve fundamental revisions to an indicator resulting from new thinking precipitated by a learning process. This is not surprising for very new indicators such as those for social impact investing. Where indicators are being relied upon by external bodies, as with GAVI's use of the immunization coverage indicators, or where the indicators have been built into standard operating procedures and country expectations, a change in methodology may be disruptive. More generally, major revisions of an indicator impair comparisons over time and are costly to implement. The business model of an indicator producer may favor stability in the indicator for these reasons and to retain the confidence of users; but in some cases a rethink of the indicator or a change in country rankings might be newsworthy and draw welcome attention to the producer. Revision of indicators after strong contestation and defense of the indicators, is also likely to depend on the dynamics of a political process. This is because, as this report argues, indicators can be important interventions, with much at stake, including the stakes of winners and losers. In these respects, an indicator can operate much like a legal standard.

A checklist for producers and users of indicators is included as an appendix to the report.

1. Introduction

Indicators are ubiquitous in development policy. Corruption. Human trafficking. Obstacles to doing business. Child mortality and morbidity. Human development. State fragility. There are indicators for all of these social problems. The range of users extends from large inter-governmental organizations and national government agencies of developed or rapidly-growing economies, to international NGOs or foundations or corporations, to national and sub-national governmental agencies and news media and NGOs in developing countries, to local communities, to individuals anywhere whose decisions may be affected by information conveyed in indicators. Some of these actors produce or commission their own indicators. Many are simply consumers of indicators produced by others.

Why the growing interest in indicators? Why do many people in the field feel unease about the growing reliance on indicators, or about the influence of specific indicators? To answer these questions it is critical to understand that the mere act of producing an indicator has potentially significant effects on development outcomes. Rather than viewing indicators solely as methods of selecting, evaluating or publicizing development interventions, the production of an indicator should be viewed as an intervention in its own right.

Surprisingly little work has been done to understand the effects of producing indicators, especially in global governance and transnational contexts relevant to development. This report draws on three case studies and a review of the literature to provide a framework for understanding those potential effects.

Indicators are:

- diagnostic tools for identifying problems and needs
- instrumental measures of performance
- techniques of awareness-raising and public advocacy
- instruments of change

Indicators are interventions!

1.1. Top-down problem-solving and the pathologies and pitfalls of indicators

'Top-down' problem-solving approaches to complex social issues have a checkered history and have attracted some well-founded critiques. Much has been learnt in many decades of criticism of simple problem-formulation/solution-formulation approaches. Over-simplified and overly confident formulations of what are actually complex and poorly understood problems can lead to misguided solutions. This is especially true of approaches to problem-formulation that rely on data susceptible to manipulation or gaming by interested parties.

We also know that these formulations of problems and solutions are themselves an assertion of power over other people who are often much more directly affected by the issue than are those doing the formulating. Some such formulations have stigmatizing or disempowering

or power-aggrandizing effects. On many issues, interested actors differ sharply as to what the 'problem' is and what would be a successful intervention in relation to it. For anyone proposing an intervention it is essential to engage with these differences and seek to form shared understandings among the actors, or at least to reach some consensus on markers and measures and depictions with regard to the issue and ways forward (these insights are utilized in work on 'wicked problems' as well as in other work).

Formulations of 'problems' and 'solutions' (or, often more realistically than 'solutions', ameliorations or ways forward) must be revised as experience is learned from, perceptions change, or wider conditions change. Some issues are in great part systemic: they can only be effectively addressed through an understanding of a complex system in which they are embedded, and the continuous relations of accommodation or friction or disruption between that system and others. A single 'problem' or 'solution' may be directly or indirectly dependent on multiple complex systems, which the key actors may in turn understand differently or not really understand at all.

'Top-down' problem-solving has made extensive use of indicators. In some cases, indicators have been a significant part of 'the problem' with simple problem-oriented approaches. Indicators frame 'the problem'. Often they embody, at least implicitly, theories about what causes the problem, how those causal features can be overcome or redirected positively, who the key actors are for the problem to be addressed productively, what pathways are likely to produce better outcomes, and what would constitute success in relation to the problem from the viewpoints of key actors. Some of the critiques of top-down problem-oriented approaches are critiques of the use of indicators in such problem-solving models, and vice versa.

1.2. Indicators as interventions aimed at wicked problems

The 'wicked problems' approach to complex social issues has come to the fore as a response to many of these critiques of top-down problem solving.[1] While still open to critique, it has the potential to place emphasis on what every actor involved understands the relevant problem to be, and to incorporate the views of these actors on both the diagnosis and desirable or undesirable steps to take.

Wicked problems have several characteristics which make them challenging. Wicked problems are difficult to understand in single standard categories – for example, state failure is variously addressed by external agencies as a development problem or a security problem, and local people may see it in other ways. Wicked problems are complex: they typically have many causes, and potential responses often involve a myriad of different actors. Moreover, these problems take different forms in different places, making each version of the problem somewhat unique. Significant interventions in response to wicked problems are likely to have consequences that are not easy to anticipate. In some cases, peoples' lives will be

[1] Horst Rittel and Melvin Webber, "Dilemmas in a General Theory of Planning," 4 *Policy Sciences* 155–169 (1973).

irreversibly influenced and circumstances irrevocably changed. For all these reasons it is often unrealistic to speak of *solving* these kinds of problems. On the other hand, it is not unrealistic to speak of *addressing* them constructively. Finding the most constructive interventions is challenging. Indicators, which seek to simplify and typically are not participatory, may seem an unpromising form of intervention in such contexts. We suggest, however, that well-crafted indicators may play valuable roles.

Indicators have potential to be constructive interventions in addressing *wicked* or even *super-wicked* problems of development. The diverse actors whose efforts are needed to fruitfully address a wicked problem do not see the problem the same way or with the same parameters; most may identify with only some aspects; each may pursue wider ideals or agendas that unwittingly pull away from collective improvement; they may hold collectively incompatible views due to responding to different questions; they may not see themselves as engaged in any way in the same problem as some of the other relevant actors. Well-crafted indicators may build shared understandings and coherence, mobilize key actors in their responses, and help bring them into a common enterprise with other actors essential to addressing that problem. Indicators can be designed in ways that generate learning, and they can be built into structures for dissemination of this learning and alteration of practices to reflect learned experiences. More fundamentally, indicators can be piloted in a first-approximation formulation of the 'problem' and/or the pathways of change and/or the 'solution', then amended or radically reconsidered as new framings or new views of the problem or the pathways or the solution emerge or become dominant among the actors, or to build in feedback effects, or to respond to system-scale or inter-system impacts.

Indicators: Pathologies and Pitfalls

- *Allure of quantification*: excessive belief in veracity and accuracy of numbers
- *Uncertainty absorption*: caveats about measurement and data quality entered by technical producers but ignored by decision-makers and media
- *Simplification*: complex problems reduced to simplistic single scores
- *Distortion of priorities*: "what's measured is what matters", but what matters is often not measured or measurable
- *Gaming* of indicators, diverting effort from better paths
- *Zero-sum rankings*: there must be a loser for every winner
- *Entrenchment of ideas*: mindless reliance on indicators can stifle real debate, penalize experiments and innovation, and reward short-term quick-fixes
- *Quantitative arms race:* "it takes an indicator to beat an indicator"
- *Technocracy*: excessive power given to indicator designers and producers with political choices and cultural preferences disguised as technical issues
- *Lack of accountability or participation* in indicator production
- *Instability*: Frequent revisions of indicators reflect new priorities of power-holders rather than learning, and dis-incentivize performance improvements

Our work so far has not found examples of sophisticated portfolios of development-related indicators designed and managed to do all of this. However, practice in the production and use of indicators bears (intentionally or inadvertently) on many of these issues. We believe a great deal can be learned from study and analysis of this experience, and from further thought, experimentation and critique in these areas. Drawing on a review of the literature and three case studies, this report provides some analytical foundations and illuminating examples. But much more work is needed on these issues.

In the discussion that follows, we delineate some of the key considerations relevant in determining whether, and how, indicators may be employed constructively as interventions addressing wicked problems related to development. Two sets of considerations will be examined.

One set of considerations relates to the framing of a complex problem and the decision processes for taking and reviewing actions to address the problem. How do indicators influence the understanding of a complex problem: its parameters and components, causes and potential solutions? How do indicators influence decisions on interventions to address the problem? Do indicators help enable, or discourage, specific types of contestation (or the influence of particular actors) about framing of the problem or decisions on responses to it?

A second set of considerations relates to the large number of actors whose actions and understandings must be aligned to deal effectively with a complex social problem. We examine how production of indicators influences variations in understandings, decisions and actions across groups of actors in these multi-actor contexts. The report seeks in this respect to illuminate when and how the production and use of indicators can serve to:

- create common understandings of how to frame, analyze and respond to a particular problem;
- create social structures that can be used to facilitate common understandings and trust across a range of interactions;
- enhance the efficiency, transparency, consistency, and perceived legitimacy of decision-making;
- ensure that decisions about both the framing of problems and possible steps to improvement are open to contestation by a broad range of interested parties; and
- enhance the (positive) impact of policy and programmatic interventions on the behavior of actors involved (including donors) in ways that improve responses to social problems.

1.3. Overview of the report

The remaining sections of this report are as follows. Section 2 offers a definition of the concept of an indicator and provides an overview of how indicators are produced and used. This section also compares and contrasts indicators with alternative ways of intervening in the process of addressing social problems. Section 3 sets out several general propositions

about the effects of using indicators. It identifies two channels through which indicators influence responses to complex social problems, namely, by providing information and by stimulating engagement.

Sections 4, 5, 6 and 7 respectively draw upon both the theoretical literature and our case studies to examine the roles that indicators play at each stage in the process of addressing social problems: framing, action, contestation, and learning/revision. A checklist for producers and users of indicators is included as an appendix to the report.

2. Indicators and their alternatives

2.1. What is an indicator?

The term 'indicator' can be defined as follows:

> *An indicator is a named collection of rank-ordered data that purports to represent the past or projected performance of different units. The data are generated through a process that simplifies raw data about a complex social phenomenon. The data, in this simplified and processed form, are capable of being used to compare particular units of analysis (such as countries or institutions or corporations), synchronically or over time, and to evaluate their performance by reference to one or more standards.*[2]

This report is concerned primarily with *social indicators*. These are indicators that claim to capture information, in the problem-oriented perspective outlined in the Introduction, about the current or anticipated extent of *social problems*. For our purposes a social problem is any characteristic of a society that can potentially be improved through deliberate human action. These problems can be defined more or less broadly—for example, the prevalence of infectious diseases as opposed to the prevalence of HIV. Our main focus is on complex problems involving many actors, and in particular, complex problems that affect the most vulnerable members of developing societies. Some of these are 'wicked problems' in the sense that different actors whose behavior or attitudes or support are important to addressing the issue, have sharply different views about what the 'problem' is, and about what would be a 'solution' or at least a success in relation to it; they may also have different views about what 'causes' the problem and what the causal pathways are toward what they perceive as solutions and non-solutions. Further challenges arise where the actors have different opinions, or different understandings, of systemic considerations bearing on the issue.

There are many examples of indicators of complex social problems. These indicators vary in the identity of the producers – international or local, non-profit or for-profit, individual or a collaborative – the identity of users, the category of uses, and the kind of impact the indicator has.

The Political Risk Index, an indicator based on 12 political and social variables that is produced by Political Risk Services, a for-profit entity, is designed to be used by investors in deciding where to invest and what risk premiums to charge. In addition, one or more Political Risk Services sub-indicators form a component of each of the World Bank Institute's governance indicators.

Source: www.prsgroup.com/icrg.aspx

[2] Kevin E. Davis, Sally Merry and Benedict Kingsbury, "Indicators as a Technology of Global Governance," 46:1 *Law and Society Review*.

The Africa Media Barometer, measures press freedom through assessments by media practitioners and civil society in each country using 45 indicators. The media project of the Friedrich-Ebert-Stiftung (FES) in Africa orchestrates this and uses the information to raise public awareness for media protocols and declarations signed by governments but not implemented through national law, raise awareness among NGOs of the importance of media issues for "good governance", and provide the basis for a reform agenda to media NGOs.

Some indicators are used to monitor and evaluate the extent of problems and the success of interventions. Many of these are issued annually, so anticipation of the next publication can influence behavior of interested stakeholders. Immunization coverage indicators produced by the World Health Organization and UNICEF (discussed in detail below) allow for diagnosis of national-level under-immunization as well as track countries' performance over time. Producers of the Failed States Index (also discussed in detail below) promote it as providing a "snapshot" of a country situation, in order to draw attention to social, economic, and political trends in that country. The Trafficking in Persons Report, issued by the U.S. Department of State, allows the U.S. Congress to monitor states' compliance with anti-trafficking standards and note progress (or lack thereof) over time by moving countries among ranked tiers depending on their performance. Human rights indicators are in the process of being developed by the Office of the United Nations High Commissioner for Human Rights (OHCHR). If this project proceeds as planned, which is not certain, the 12 indicators will represent the right to life, the right to adequate food, the right to judicial review of detention and the right to health, as well as cross-cutting indicators on violence against women and non-discrimination. The OHCHR indicators are separated into structural indicators that reflect the ratification of legal instruments, process indicators that assess state policies to promote the right, and outcome indicators to assess realization of the right. The indicators are not meant to replace qualitative reports: they are only one part of the reporting and monitoring process. Their primary intended use is to compare a country's performance over time rather than against that of other countries.

Indicators can also have a community mobilizing effect. They can bring attention to a problem, drawing in actors from seemingly disparate areas of interest and expertise, and allow those actors to develop new and shared understanding of the problems, learn about causal relationships between actions and outcomes, share information about past or anticipated actions, and cooperate across a range of activities. For example, the International Institute for Sustainable Development (IISD) in Canada

Community Indicators Victoria aim "to support the development and use of local community wellbeing indicators in Victoria, Australia, with the purpose of improving citizen engagement, community planning and policy making" by providing a community wellbeing indicator framework, measuring everything from crime to household waste recycling to education levels to work-life balance. The indicators are intended to inform local policy makers' choices of initiatives, but they are also a means for social organizing and community education.

Source: http://www.communityindicators.net.au

supports local production of indicators by First Nations communities. IISD's support of local projects is designed to empower individuals and groups to develop their own indicators as a means of social mobilization. This mobilization is somewhat aligned through inter-jurisdictional collaborations that enable iterative experimentalist learning. Comparability is also facilitated by use of standard 'community indicator resource packs' and other coordinating methods.

In addition to mobilizing stakeholders for a particular intervention, indicators can provoke action - including national policies and government action -- qand guide allocation of funds. Following mediocre results on the Programme of International Student Assessment (PISA), an influential set of indicators of the achievements of schoolchildren promulgated by the OECD, Germany initiated substantial educational reform, and the federal government set up an educational program with a budget of 4 billion Euros.

The Freedom in the World Index, produced by Freedom House (a U.S. NGO), as well as the Global Corruption Barometer produced by Transparency International (a multinational NGO), are used by the Millennium Challenge Corporation, a U.S. foreign aid agency, as one of the six indicators that measure the extent to which countries are "ruling justly". Ruling justly is one of three criteria countries must satisfy in order to be eligible for foreign aid from the MCC. WHO/UNICEF immunization coverage indicators are used as performance indicators for development assistance by the World Bank and as one of criteria for determining a state's eligibility for funding from the MCC.

The Country Policy and Institutional Assessment (CPIA), an indicator of the quality of countries' policies and institutions promulgated by the International Development Association (IDA) – the branch of the World Bank responsible for providing financing on a concessional basis – is an important factor in determining the IDA's allocation of aid across eligible countries (other factors include population, Gross National Income per capita and the recent performance of World Bank projects in the country). The CPIA assesses countries on the basis of 16 criteria, which fall into four subcategories: economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions. The final category, public sector management, is also known as "governance." For each of the 16 criteria, a country is given a score ranging from 1 (low, or weak) to 6 (high, or strong).

Source: World Bank, 2010 CPIA questionnaire

2.2. How indicators are produced

Production of an indicator can be schematized as a five-stage process.

1) *Data collection*. Data can be obtained from all kinds of sources, including: interviews; direct observation of behavior; public statistical agencies; newspapers, and other media; or surveys, either of firms or individuals, experts or lay people.

2) *Processing* of the data in accordance with a specific methodology. It is not uncommon for processing to involve aggregation of data from multiple sources. It can also involve filtering that excludes certain data: for example, excluding outliers or other data deemed to be unreliable or irrelevant. Sometimes data are filtered out and replaced with statistics, such as means or standard deviations, meant to convey similar information. In still other cases missing data is filled in with values estimated from existing data. These methods of processing are frequently characterized—or perceived—as being consistent with 'scientific' methods. It may happen, however, that a scientific justification (e.g. 'incomplete data') is given for an essentially political decision not to include an adverse entry for a particular country.
3) *Promulgation* of the indicator, which involves one or more specific actors attaching their name and imprimatur to the indicator.
4) *Distribution* of the indicator. Distribution techniques vary in the extent to which they permit the user to search for and retrieve specific data (such as the scores on component indicators forspecifiedcountries in a specified time period).
5) *Publicizing* the creation or expansion of an indicator through press releases, blog postings and the like. (This is not essential to the production of an indicator. The business model of the promulgator may affect the intensity of its efforts to publicize the indicator.)

The production of indicators is often a collective process. In many cases promulgators attach their names to indicators to which a number of other actors have contributed, like the 'manufacturer' of an athletic shoe, whose main contribution is to lend its brand name to the collective product of a global supply chain. (Davis, Kingsbury & Merry, 2012) For example, PISA reports and rankings are promulgated by the OECD but are actually prepared by an Australian consultancy under a contract with the OECD. The data collection process can also involve a large network of independent actors ranging from international agencies, to national statistical agencies, to local and national NGOs, to villages and local communities, to academic researchers, to for-profit polling companies. Meanwhile the methods used to process the raw data typically rely on contributions from some segment of the scientific community. Finally, publicizing the indicator often involves close interaction between for-profit media outlets and all of the actors listed above.

There can be various sorts of legal and economic relationships among the contributors to an indicator. For instance, many social indicators are distributed free of charge or at prices that are well below their cost of production. Sometimes this happens because their promulgators are international organizations or not-for-profit entities whose mandates allow them to engage in activities that do not generate economic returns.

In other cases the promulgators are profit-seeking entities who receive indirect economic benefits, such as free publicity, from the use of their indicators.

2.3. Alternatives to indicators

Any organization considering whether to intervene in the process of addressing a social problem by producing or using indicators should consider how those indicators compare to alternative interventions.

In many cases the set of relevant alternatives will include alternative indicators. In addition, there are alternatives to indicators that involve other ways of presenting data about the extent of the social problem:

- Non-numerical data such as photographs, videos, songs, and slogans
- Combinations of non-numerical data and numerical data: for example, the U.S. State Department's reports on human trafficking include numerical country rankings, and a set of vignettes about particular individual victims expressed in narrative form
- Legal processes and verdicts
- Qualitative assessments and reports
- Systematized searchable raw data that the producer does not transform into an indicator with rankings etc. but which allow users either to build their own indicator or simply use the raw data

Interventions that involve indicators should also be compared to interventions that go beyond merely providing data. If the ultimate purpose of intervention is to effect social change then indicators should be compared to familiar interventions, such as direct provision of money, goods or services; capacity-building; regulatory change; litigation; etc.

It is also important to consider potential interactions between indicators and other interventions. Depending on the situation, indicators may be complements to, competitors with, substitutes for, or substituted by, other interventions. For example, where a legal standard is framed in very general terms, indicators and other quantitative measures can be used to help give precision and content necessary for routine concrete application. In these situations indicators may even be built into legal structures: for example, tax deductions may only be available for donations to charities that achieve a certain "quality" rating in terms of social impact or other criteria such as organizational efficiency, and such ratings may be determined by indicators. Credit ratings already play this kind of role in legal governance, both nationally and under the Basel capital adequacy accords.

2.4. Case Studies

In preparing this report our research team has conducted in-depth case studies of three specific indicators[3]:

[3] Modified versions of these studies appear in Kevin Davis, Angelina Fisher, Benedict Kingsbury and Sally Merry (eds), *Governance by Indicators: Global Power Through Quantification and Rankings* (Oxford University Press, 2012, forthcoming).

- Immunization coverage indicators produced by the World Health Organization and UNICEF, which measure the number of children vaccinated as a proportion of the target population (Angelina Fisher)
- The Global Impact Investment Rating System, GIIRS being developed by B Lab, a non-profit based in Pennsylvania to "assess the social and environmental impact (but not the financial performance) of companies and funds using a ratings approach analogous to Morningstar investment rankings or S&P credit risk ratings" (Sarah Dadush)
- The Failed State Index, an index of state failure produced annually by the US Fund for Peace (USFfP), and published in *Foreign Policy* (Nehal Bhuta).

In addition to focusing on the specific indicator, each of the three cases also examined the roles that other indicators, as well as alternatives to indicators, play in addressing the problem that defines the case. The case studies, which draw on document and literature reviews as well as numerous interviews, are summarized below.

1. WHO/UNICEF Immunization Coverage Indicators

WHO/UNICEF immunization coverage indicators measure the number of children vaccinated as a proportion of the target population. These indicators were initially intended to evaluate progress towards target immunization rates at a country level. Immunization coverage indicators are generated using a combination of administrative data, usually reported from health facilities based on the country's routine monitoring system, and surveys. National coverage data are reported annually to WHO and UNICEF on the Joint Reporting Form on Immunization (JRF). WHO then analyzes the data, arrives at national estimates, and relays the estimates back to countries' Ministries of Health so any concerns from country representatives can be voiced to WHO/UNICEF and taken into consideration as the estimates are reviewed and finalized.

WHO/UNICEF immunization coverage indicators are used by international organizations to evaluate their own work and to motivate donors by demonstrating the impact of their investment in supporting immunization. Donors also use the indicators to evaluate the performance of a country's immunization delivery systems in order to determine the type of funding a country may be eligible to receive. The Global Alliance for Vaccines and Immunization, for instance, precludes countries with less than 70% DTP3 (third dose of diphtheria-tetanus-pertussis vaccine) coverage from receiving funding supporting introduction of new vaccines. The reason for this threshold appears to be driven by donors who feel that since low DTP3 coverage indicates that countries are not able to handle routine immunizations, it would be inefficient and even wasteful to supply new vaccines to such countries.

Over the years, the use of immunization coverage indicators broadened beyond evaluation of immunization systems. Ease of measurement, routine provision, and wide availability made the indicators, in particular DTP3 coverage, an appealing proxy for the quality of the country's overall health system. Development agencies, including the World Bank and the MCC, use immunization coverage as one of the determinants of funding. In addition, the

World Bank uses immunization coverage indicators to evaluate progress toward addressing poverty reduction for the purpose of determining eligibility for aid (i.e. debt relief) from international financial institutions.

Though immunization coverage may be suggestive of progress toward poverty reduction in the social sector, concerns arise that reliance on these indicators by development agencies places intense political pressure on program managers to achieve stated targets, sometimes using questionable means, and diverts resources from other poverty reduction measures that may be more directly effective.

2. State Failure: The U.S. Fund for Peace Failed State Index

The classification of a country as a fragile or failing state has implications in both the development assistance context, where a state in desperate need of aid may not meet performance-based requirements for assistance, as well as a security context where a state may warrant special scrutiny or multilateral diplomatic action. This lack of specificity complicates measurements of state failure because the concept assimilates security and development. Broadly linking security risks with underdevelopment could not only skew development priorities but also incentivize deeper and more intensive forms of intervention in the developing world.

The Failed State Index is produced annually by the US Fund for Peace (USFfP), and published in *Foreign Policy.* It ranks 177 countries based on an aggregate of their scores on 12 indicators. Each country is scored internally by the USFfP between 1 (best) and 10 (worst) for each of the twelve indicators based on a combination of qualitative and quantitative methods in a system that the USFfP has not yet disclosed. The 12 indicators for which each state receives a score are as follows:

1. Mounting Demographic Pressures
2. Massive Movement of Refugees or Internally Displaced Persons
3. Legacy of Vengeance-Seeking Group Grievance or Group Paranoia
4. Chronic and Sustained Human Flight
5. Uneven Economic Development Along Group Lines
6. Sharp and/or Severe Economic Decline
7. Criminalization and/or Delegitimization of the State
8. Progressive Deterioration of Public Services
9. Suspension of the Rule of Law and Widespread Violation of Human Rights
10. Security Apparatus Operates as a State within a State
11. Rise of Factionalized Elites
12. Intervention of Other States or External Political Actors.

Due to the numerous points in the process at which the basis for aggregation of different kinds of data is not verifiable, the basis for comparability between the indicator levels for each country is uncertain. Comparability is further complicated by what seems to be an inherently relative scoring standard, in which a "10" for a particular state may be determined by reference to specific historical examples representing a nadir for state failure for that state. As such, the ordinal ranking – which generates so much attention for the index – may not be useful as a measure of *degrees* of state failure.

There is no evidence that the Index has a determinative effect in decision-making, and the creators of the Index promote it more as providing a "snapshot" of a country situation in order to draw attention to social, economic, and political trends in that country. It appears, however, that the ranking also spurs engagement from some states, and that the Washington-based diplomats of ranked countries pay some attention to the placement of their state.

Despite unresolved questions about the nature and effects of these new indicators, there continues to be interest in the creation of indexes of state fragility or failure. Multilateral development banks are increasingly involved in states affected by conflict, and seek benchmarks to establish eligibility for special programming or for the relaxation of performance-based evaluations. Development assistance agencies of first world states are interested because "post-conflict reconstruction" has formed a growing component of their work, leading to an operational convergence between development assistance, institution-building, and attempts to address "root causes" of violence and civil conflict. The interest of development agencies has also coincided with the high level of interest from military and intelligence agencies in the security threats posed by "ungoverned spaces". These actors are concerned to find ways to preempt severe disorder and state collapse, through early warning, civilian response, stabilization/reconstruction, and targeted aid/development assistance.

3. Impact Investment: The Global Impact Investing Rating System (GIIRS)

Impact investment is defined as "actively placing capital in businesses and funds that generate social and/or environmental goods and at least return nominal principal to the investor."

Many social impact indicators and ratings systems exist in the market. One of these indicators is the Global Impact Investing Rating System, GIIRS, which was created in 2009 by B Lab, a U.S. based non-profit. GIIRS is a ratings system made up of a set of indicators that seek to capture the social and environmental impact associated with investing in a particular firm or investment fund. GIIRS, which is currently under development, is intended to serve a purpose analogous to that of the S&P Ratings system.

GIIRS ratings take the form of stars, ranging from 1 to 5, and percentage scores based on a possible maximum score for a similarly placed company. Companies receive both an overall rating and ratings in 15 sub-categories. GIIRS also provides key performance indicators that vary according to the type of company. For investment funds, GIIRS provides a rating that is based on the aggregated and weighted ratings of a given fund's portfolio companies.

B Lab is currently working with the Standards Advisory Council, as well as with entrepreneurs, investors, and policymakers regarding the content and weightings of the ratings system. This process included a private beta launch of the Impact Investing Rating System in September 2010 where the Pioneers GIIRS funds and the companies in which they have invested test the tool and provide feedback. This was followed by a public beta release

of the tool in January 2011, which was open to anyone and was equipped with feedback mechanisms.

GIIRS' intended users are larger investors (i.e. institutions, high net worth individuals, foundations, donor-advised funds), rather than consumers and other stakeholders. While the fee structure was still being finalized at the time of our case study, the thinking was that basic information on the website, including star and percentage ratings for companies or investment funds would be available free of charge, while details on the data used to produce those ratings (including the social/environmental effects of adding a rated company to an investment portfolio) would be available only to paying subscribers. Additionally, companies and investment funds would pay to be rated.

Four main categories of challenges are faced by GIIRS. A first challenge relates to whether GIIRS will be able to provide accurate measurements of the social impact of the firms receiving a GIIRS rating. A second challenge will be to establish the legitimacy of GIIRS given the limited role that beneficiaries play in its governance structure. A third challenge will be to distinguish GIIRS from competing social impact indicators. Finally, there are grounds for concern about whether using indicators to enhance access to capital is the most effective way to intervene if the goal is to help businesses maximize their social impact.

3. Roles for indicators in addressing social problems

This chapter focuses on the possible roles of indicators in interventions formulated on the basis of problem-solution approaches that incorporate insights from work on wicked problems and complex systems.

As noted in the Introduction, problem-solution approaches can themselves be problematic. As well as other difficulties, the assertion of power that these approaches involve can raise issues of justice, voice, responsibility, accountability, and political morality. It can provoke objections and contestation that has practical consequences. It may shape the identity of or stigmatize groups or persons labeled as being part of the problem. In India, for example, opposition to, and remedial processes in the name of, groups adversely affected by large infrastructure projects such as the Sardar Sarovar Narmada dam produced social categories of 'project-affected people'. The designation of certain states as being, or being on a continuum toward, a 'failed state' is not necessarily one that fosters imaginative leadership or an influx of private foreign investment. While recognizing these reasons for care, we proceed on the assumption that approaches compatible with the 'wicked problems' concept are likely to be influential, and that they can potentially be fruitful in some development contexts.

3.1. The process of addressing social problems: framing, action, contestation, updating

The process of addressing a social problem can be modeled in four stages: framing, action, contestation, and learning/revision.

1. Framing

Framing a particular problem involves deciding what the problem is – what is within its parameters, and what is marginal or only obliquely connected or lies outside and is excluded. Framing extends to determining which outcomes represent improvements over other outcomes and specifying a set of available and feasible actions for all of the relevant actors.

Where a single actor has only to consider her/his/its own viewpoint and action, framing the problem and assessing possible action could conceivably be based purely on the single actor's own convictions. Typically though an actor framing a problem will be influenced by externally-provided information about what actions are possible and which states of the world represent improvements over others. Where multiple actors are involved – as is inevitable in relation to complex social problems – the actors may disagree vehemently with each other on these issues. This lack of a shared view about the nature and causes of the

problem, and about who the key actors are and what they should be expected to do, can make efforts to understand and solve such problems ineffective.

2. Action

In theory, once a complex social problem has been adequately framed, actors should rationally select from among the possible actions one or more that lead to the optimal outcome. In practice, the move from understanding to action is more complicated than that.

To begin with, actors generally do not know the current or future outcomes of their actions with certainty, all the more so in fast-changing environments. This uncertainty can stem from lack of knowledge about either the society in question or, given what is known about the society, causal relationships between actions and outcomes. A second complication is that outcomes are typically affected by the actions of many different actors, each trying to address distinct problems, social and otherwise. Moreover, many of those actors behave strategically, that is to say they alter their actions in anticipation of, or in response to, one another's actions. For all these reasons in addressing a problem it helps to know as much about the context as possible, including how other actors will frame the problem and what actions they have taken or will take. It is also useful to foster cooperative behavior, not only to make it easier for actors to predict one another's behavior, but also in order to avoid situations in which strategic behavior leads to perverse outcomes.

3. Contestation

Decisions or actions taken in an effort to address a problem may provoke criticism, resistance, and demands or proposals for a different approach or a different view of the key problem. This kind of contestation should be regarded as a normal part of the process.

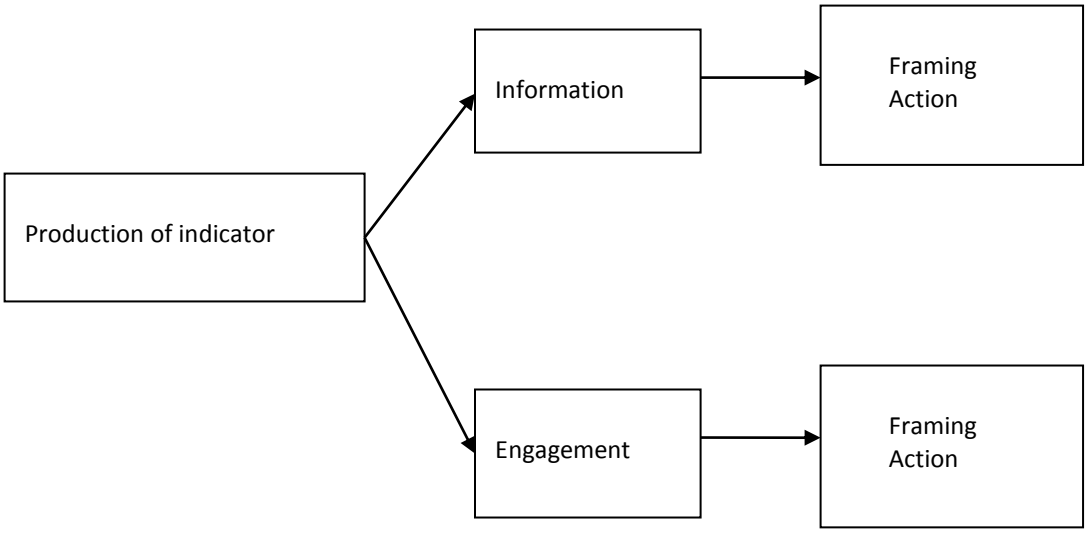
4. Learning and Revision

It is rarely the case that a single round of actions can successfully address a social problem. In addition, with the passage of time the nature of the initial problem, or at least common understandings of it, may change. For all these reasons addressing a social problem should be thought of as an iterated and collective process that involves repeated instances of reframing the initial problem and revisiting previous decisions and actions, often in response to contestation and with an eye to recent events and changes in knowledge.

3.2. Two roles for indicators

There are two distinct channels through which the production of indicators can influence the process of addressing a social problem: an informational channel and a socio-political channel.

Figure 1



The informational role of indicators

The most straightforward channel through which indicators can affect the process of addressing social problems is by serving as sources of information. In other words, indicators have the potential to prompt changes in beliefs about the world, including beliefs about the moral implications of various social outcomes, about relationships between actions and outcomes, and about the beliefs or actions of other actors. By altering the beliefs actors draw upon when deciding how to address social problems indicators can in turn affect the decision-making process, actions and, eventually, outcomes. Moreover, over time any given indicator may crowd out other sources of information; or, alternatively, it may stimulate the development of alternative information sources.

It is important to stress, however, that these effects are not automatic. On their own, indicators are simply *data*, stimuli that may or may not affect the people who receive them. In a world awash in data it would be wrong to assume that all data are treated as *information*, that is to say, the kind of stimuli that affect people's beliefs about the world.

In theory, the extent to which any given indicator becomes an important source of information –in other words, the demand for indicators – should depend on the perceived benefits and costs of relying on that indicator as compared to either existing knowledge or other potential sources of information. The most obvious benefits are reflected in the extent to which people improve the outcomes of their actions by relying on the indicator when deciding how to act. These benefits will depend in turn on, among other things, the indicator's reliability and validity. Costs will be the costs of accessing and analyzing the indicator. As with other products though, factors besides measurable costs and benefits are

likely to affect the demand for indicators. Those factors might include popularity, salience and the extent to which the indicators provide reassurance by confirming existing beliefs.

The socio-political role of indicators

The second channel through which production of indicators can influence the process of addressing social problems is through the social and political engagement they foster.

Many indicators are produced by networks of individuals and organizations. In addition, once indicators have been produced it is not uncommon for other actors to respond by criticizing or commenting on either their methods of production or their use in particular decision-making contexts. Those actors also form networks of their own. Both these kinds of networks often span multiple societies.

Once they have come into existence these networks permit many sorts of information to flow between their members, not just information related to the task that prompted the creation of the network. As a result they provide opportunities for people to develop new understandings of how to frame problems, to learn about causal relationships between actions and outcomes, and to share information about past or anticipated actions. Social networks can also foster cooperation across a range of actions. At the same time, use of indicators might drive other networks into desuetude by drying up flows of information that sustain them.

The Human Development Index of the United Nations Development Programme (UNDP) and the Doing Business indicators produced by the World Bank Group have each had such effects in creating and intensifying links among groups and individuals interested in the particular approach to development that they embody. Each has spawned further sets of indicators produced by more specialist sub-groups nurtured through the network, such as the Gender-related Development Index and the Multi-Dimensional Poverty Index in the UNDP, and the Investing across Borders index in the World Bank Group.

Going beyond simple rational cooperation or other rational interest-based behavior of unchanging actors, indicators can have more far-reaching effects in changing some of the key activities, and even changing the identities, of certain actors. The shift to use of indicators may also have effects on the social status of different actors. These may be connected with struggles for membership and influence and role and even identity between different organizations and networks.

4. Indicators and the framing of social problems

4.1. How indicators frame problems

Indicators inescapably frame the problems they measure, albeit implicitly. The label attached to an indicator amounts to a claim that the indicator provides a meaningful representation of the extent of the problem referred to in the label. In other words, the indicator embodies claims about the existence and nature of a problem, as well as how to measure it. So for example, an indicator of state failure embodies claims about whether there is such a thing as a failed state, what it means to say that a state is more or less of a failure, and how to measure the extent of state failure. Similarly, social impact indicators identify the social impact of business as a social concern, and embody claims about what count as more or less positive social impacts as well as how to measure them. Immunization coverage indicators can be analyzed in similar terms.

Indicators can influence how people frame problems in several different ways.

First and most obvious is influence through the informational channel. People who are exposed to the indicator may come to accept the claims about framing that it embodies. For instance, people who are introduced to a failed state indicator may come to accept the idea that state failure is a problem and that it is exemplified by the states that rank highly on the failed state index. Moreover, if they become familiar with the methodology underlying the indicator they may come to accept it as the 'standard' or 'appropriate' way to measure state failure.

Second, framings can be influenced as the indicator becomes 'internalized' within an organization using the indicator, through repeated use. The indicator may be built into protocols, reports, staff training etc. The indicator can have this influence through use even at the same time that it is routinely acknowledged internally to have limitations.

Third, producing an indicator can influence how problems are framed by triggering processes of deliberation and contestation. The process of designing and producing an indicator may involve disparate actors who initially frame the problem under consideration quite differently. The process of deliberating over how to construct the indicator is likely to change the views of at least some of the participants. Similarly, disseminating the indicator to a broader audience will often prompt objections to the framing embodied in the indicator. The ensuing debates might lead the participants to reconsider how they frame the problem at issue.

For example, initial phases of testing the GIIRS social impact investment indicators forced the producers to confront the dilemma of firms based in developing countries with limited control over the social performance of their supply chain. What if a firm only has access to a single supplier for a primary product and that supplier has a huge carbon footprint or

terrible labor practices? Should this kind of supply chain issue affect the rating of a firm in all countries in the same way? What if the firms in some countries systematically have less control over the social performance of their supply chains? However the issue is resolved, the process of engaging with investors facing this dilemma forced the producers of the rating scheme to confront the nuances and complexities involved in maximizing the social and environmental impact of business on a global basis.

4.2. Indicators: validity and measurement error

Some indicators frame problems better than others. In other words, some indicators provide more accurate information than others about the problems they purport to measure; more accurate indicators provide better understandings of the problem than less accurate ones.

There are at least two well-recognized sources of inaccuracy. First, indicators can lack *validity*, meaning, essentially, that they attach misleading labels to data. None of the indicators we examined was clearly invalid. However, the validity of the social impact investment indicator we studied, GIIRS, remains to be determined. While GIIRS was still under development at the time of this study, one aspect of the proposed design raises such an issue: rather than measuring achievement of positive social impact, practical considerations may force GIIRS to measure *efforts* to generate positive impacts. For example, with respect to employee engagement indicators, rather than asking about the substance of employee satisfaction, GIIRS may ask only whether there is a formal method by which employees can raise complaints. While an indicator derived from such data may be a valid measure of effort, it may not be a valid measure of impact.

Whether or not an indicator is valid depends upon the theoretical and empirical connections between the indicator and the phenomenon it claims to measure. For example, if effort to maximize social impact is a good proxy for impact then the GIIRS approach is valid. Moreover, the validity of an indicator can change as the actors being measured change their behavior over time, including efforts to try to game the indicator. For example, even if effort is currently a good proxy for social impact, it may not be once firms become concerned about boosting their social impact ratings.

A second source of inaccuracy is *measurement error*, meaning that there is a relatively high probability that if the indicator were produced again it would provide different data. Concern about this kind of inaccuracy has dogged the WHO's immunization indicators. A study published in 2008 found stark differences between WHO indicators, country-reported coverage rates, and coverage rates produced by surveys: 1990 WHO indicators estimated global DTP3 coverage at 75%, official reports at 83%, survey-based estimates at 65%; from 1999 to 2006, WHO estimates showed an 8% increase (from 71% to 79%), officially reported estimates showed a 9% increase in DTP3 coverage (from 81% to 90%), and surveys showing an increase of 4.9% (Lim et al. 2008). The study also found that in 2006, compared with survey-based estimates, official reports overestimated the number of

children immunized with DTP3 by 20.2 million and WHO overestimated the number by 6.1 million (Lim et al. 2008).

A variety of factors can cause an indicator to be unreliable. The most obvious problem is unreliable data. Collecting reliable social data is difficult, especially in disadvantaged societies: human memories and powers of observation are fallible; the most disadvantaged people may be the most difficult to contact or observe; informants may be reluctant to speak candidly to outsiders; linguistic differences give rise to misunderstandings; written records may not be kept securely; public officials or other actors might misreport data to serve their own ends; etc. Measurement error can also be introduced in the course of processing raw data to construct an indicator, for example, in the process of imputing values for missing data.

4.3. Discrepancies in framing between the indicator and the gold standard for the problem

The possible relations between the way a particular indicator frames a problem, and the way a problem ought to be framed from the subjective standpoint of one observer or from the more objective standpoint of an existing 'gold standard', can be mapped two-dimensionally. Five different relations are possible, in this view.

The indicator may, in relation to the other (subjective or gold-standard) framing, frame the problem:

- Identically
- More narrowly
- More broadly (but fully subsuming the problem)
- In an overlapping way (subsuming part of the problem and also other separate issues)
- In an irrelevant way (addressing only separate issues).

These five relations will be the subject of brief comment in this section.

Indicators are simplifications. It is thus extremely difficult to produce indicators that frame complex 'wicked problems' in a way that is *identical* to an independent 'gold-standard' framing of the problem, especially if the gold-standard framing is not quantitative.

Indicators that are framed *narrowly* can begin to define what they only incompletely proxy. The Human Development Index, for example, becomes equated with "human development" just as the GDP/capita becomes "economic development." Nonetheless, successful indicators typically embody narrow framings. Several indicators may be used, each of which deploys a different narrow framing (and these different narrow framings may overlap). For example, an immunization coverage indicator based on the number of children vaccinated frames the problem of under-immunization in a different way from an indicator that measures incidence of vaccine-preventable diseases among children.

An indicator may frame the problem more *broadly* than needed, in a way that captures

several narrower framings. This can occur particularly with composite (or mash-up) indicators. Take for example indicators of state failure. Most people would agree that failed states are states experiencing exceptionally poor governance. Some people view that as a problem because it implies that the state is unable to promote the development of its population. Others view poor governance as a problem because poorly governed countries are at risk of generating local, regional or global security threats.

These contrasting views of the nature of the problem are also associated with contrasting views of the causes of and appropriate responses to state failure. In one view, state failure is a form of institutional underdevelopment and the appropriate responses are various kinds of development assistance and capacity-building. In another view, the causes of the problem are more likely to include legacies of internal conflict and lists of the appropriate responses are more likely to include military intervention. These are two dramatically different ways of framing the problem of state failure. In principle, indicators of state failure could embody one framing of the problem or another. Many, however, obscure the issue. However, some failed state indicators incorporate data on both economic development and risks posed to other states, and frame the problem of state failure so broadly as to conflate the two narrower understandings. Excessive breadth may undermine the indicator's value as an accurate representation of a problem and may limit its value in tracing the effects of interventions on the extent of a problem.

Indicators may also address part of a problem but at the same time *incorporate non-relevant material*. Mis-framings can cause or reinforce a mis-diagnosis of the problem and its causes, as well as mis-design of optimal interventions. In practice, however, it is often impossible to avoid some mis-framing, especially where indicators rely on data compiled for other purposes.

Some indicators purport to frame a problem, but in fact do not really address at all the problem they purport to address. This can readily occur where the data used is only *spuriously relevant* to the real problem.

4.4. Indicators that frame problems controversially

Frequently the way in which an indicator frames a problem is not clearly wrong (it does not conflict with an established gold standard), but is controversial. This happens when a problem can be framed in multiple ways – as is typically the case with 'wicked problems' – meaning that there will be room to contest any particular framing. Debate about the nature of a problem naturally translates into debate over valid ways to measure it.

Immunization coverage indicators are a case in point. Some organizations use these indicators as measures of the overall performance of health systems. For instance, the World Bank has used scores on WHO/UNICEF immunization coverage indicators as targets that must be achieved by countries seeking disbursement of loans designed to improve public health programs. Some actors claim that specific characteristics of immunization coverage indicators – ease of measurement, routine provision, wide availability, and strong impact on child morbidity, child mortality and permanent disability – make them a good measure of the overall quality of a country's health system. But these claims are contestable. In particular, the precise association between immunization and child mortality is contestable because of disagreements over which diseases qualify as vaccine preventable diseases. If the leading causes of child mortality – pneumonia and diarrhea – are not considered vaccine-preventable diseases, then immunization coverage becomes a much less valid measure of a health system's performance.

Validity of measures of Soviet Immunization

The former Soviet Union did not observe the 12-month timetable of vaccination used by WHO/UNICEF. As a result, even though almost every child was eventually immunized, many were immunized after 12 months of age. The data calculated nationally reported high coverage levels (which reflected the reality that virtually all children had been vaccinated) but the WHO/UNICEF estimates were low because many children were vaccinated after 12 months of age. It is debatable which measures best represented the extent of 'problematic' under-immunization. The Soviets' postponement of vaccination reportedly was not because they had a poor immunization delivery system, but because, relative to other countries they believed in a large number of counter-indications to immunization and so wouldn't immunize a child who exhibited some characteristic that would have gone unremarked in other countries.

Social impact investment indicators also provide fodder for debates about validity. Broadly speaking, the problem these indicators help to address is failure on the part of individual for-profit firms to maximize the positive social or environmental impacts of their operations. This broad frame captures many more narrowly defined problems, corresponding to different ideas of what counts as a 'positive impact' or how different impacts should be valued and prioritized. For some, the problem may be the environmental impact of some kinds of smallholder agricultural enterprises. Others may be most concerned about increasing employment and empowerment of women or specific ethnic minorities, while others may focus primarily on working conditions. Some people may be most concerned about impacts felt in the United States, while others may be most concerned about impacts felt in a particular developing country, or in developing countries in general. There is ample room for debate about whether any given social impact indicator gives appropriate weight to firm performance along each of these dimensions.

4.5. Do indicators promote common understandings?

The case studies in this report suggest that common understandings have been promoted by the WHO/UNICEF's processes in production of immunization coverage indicators. Some of the uses of these indicators (for example, in HIPC debt relief decisions) may not, however, either reflect or have produced such shared understandings. The WHO's short-lived attempts to rank countries' health systems through an indicator seem not to have been based on or to have produced shared understandings. As regards the GIIRS impact investing indicators, the substantial groundwork done by the promulgators may produce some shared understandings among certain categories of investors, certainly in North America and perhaps also in the North Atlantic more broadly, but the degree of commonality established with companies seeking to attract investment, or with specialized NGOs, or with government regulators, is not yet clear.

Even if the production of an indicator influences how people frame problems there is no guarantee that it will influence them in the same direction. The fact that indicators bring people together does not necessarily mean that indicators will lead them to agree. In other words, it is unclear whether production of indicators will lead to common understandings of social problems.

If each member of a group of users treats an indicator as just one of several sources of information about how to frame a problem, the group may end up with quite divergent understandings of the problem. For example, an analyst who refers to a particular set of state failure indicators together with other information may form a different understanding of the problem from an analyst who relies on a slightly different set of information; and neither of them may frame the problem in the same way as the promulgators of the indicators. One analyst may decide that the key feature of the problem is the risk of internal conflict. The other may decide that the crux of the matter is institutional underdevelopment. Meanwhile, the promulgators might be focused on the risk of insecurity spilling over into foreign territory.

Indicators that frame problems broadly may also serve to foster or reinforce divergent understandings. So for example, two different analysts relying on the same state failure indicator could easily arrive at very different understandings of the problem. One might frame the problem in terms of underdevelopment while the other views it as a security problem.

5. How indicators influence action

Indicators can do more than just influence understandings of problems, they can also influence action. Alerting people to the existence of a problem, helping them to understand its magnitude, and pointing them toward means of addressing it might stimulate at least some people to action and shape the actions they undertake. In addition, becoming involved in the process of creating or contesting an indicator can mobilize actors to focus on and address a problem.

The scope for deliberately producing indicators to influence action is large. Addressing complex social problems successfully can require many different kinds of action: action by intergovernmental and governmental leaders and agencies, action by donors and voters, action by public officials, action by for-profit corporations, actions by volunteers and not-for-profit organizations, alliances, or networks; and, usually of central importance, action by members of affected communities. Indicators have the potential to influence all of these actors.

5.1. Which indicators influence action?

Examples of influential indicators

There are certainly cases in which it is clear that indicators are influential. The most prominent examples we have studied are WHO/UNICEF immunization coverage indicators, which are used by donors to decide how to allocate funding. GAVI, for example, uses a DTP3 coverage threshold to determine the type of funding available to a country. Under GAVI's rules, countries with less than 70% DTP3 coverage are *not* eligible for funding to support introduction of new vaccines such as Haemophilius influenza type-b, pneumococcal vaccine, and rotavirus vaccine. The rationale here is that immunization coverage indicators are measures of the performance of immunization delivery systems.

In many cases decision makers appear to refer to indicators but it remains unclear whether or how the indicator has influenced subsequent decisions or actions. For instance, the creators of the Failed State Index claim that it has been referred to in foreign aid allocations. A preliminary review of documents in the USAID public database found 12 USAID country and policy reports since 2006 that refer to the FSI as part of a country-specific analysis, or as part of a discussion of a Post-conflict Assessment Template. Some of these documents refer specifically to CAST Indicators and sub-components of the FSI when drawing conclusions about the risks of political instability in particular countries. The creators of the FSI also claim that the UK Department for International Development, the Australian aid agency AusAID, and the Millennium Challenge Corporation all have used the Index as part of evaluation of aid outcomes. The significance of these uses is not clear in research

undertaken so far and no evidence has been found that the Index has had a determinative effect on decision-making in this context.

Examples of indicators that are not influential

Our case studies also provide insight into the kinds of indicators that are *not* influential. Our case study of the WHO/UNICEF indicators suggests that indicators are rejected – particularly by users needing to make operational decisions or deep-seated causal analyses – when they aggregate too much data or frame problems too broadly to be useful as guides to action.

Some potential users reject highly aggregated indicators because they obscure how the extent of a social problem varies with factors such as geography, ethnicity or gender, making it difficult to set priorities in addressing the problem. A modest example of this is PAHO, which focuses on assisting countries with improving monitoring and reporting of health systems performance, and does not use WHO/UNICEF indicators very often in its work. The main reason is that the WHO/UNICEF indicators are aggregated at the national level but PAHO is often interested in disaggregated data that makes it possible to identify variations in performance across districts. National data are also unsuitable for determining the distributional impact of immunization. So for instance, PAHO sometimes uses immunization data disaggregated by ethnic group instead of the WHO/UNICEF indicators.

Other potential users reject indicators based on highly aggregated data because they frame problems more broadly than the problems potential users are equipped to address. So for example, immunization coverage data disaggregated at the sub-national level are more useful than WHO/UNICEF national data in attracting and sustaining local leaders' (community leaders, politicians) interest in immunization.

Why we know so little about which indicators are influential

The literature review conducted in the course of preparing this report presents some of the very scant evidence on the effects of indicators in transnational or global governance contexts. That there is considerable anecdata, but little robust data, is not highly surprising. It can be difficult even to determine the extent to which any given indicator is communicated to particular individuals; and even if that hurdle is overcome, determining the extent of the indicator's influence on their ultimate actions is much more challenging. The main source of difficulty is that actors often rely on many different kinds of data in the course of deciding whether and how to address problems, and they do not necessarily understand, record or disclose the elements of their decision making processes. Even among organizational actors, those with highly formalized and transparent decision-making processes are the exception rather than the rule.

The mere fact that an indicator is referred to by a person does not mean that it has any influence upon their subsequent actions. This is particularly true of large organizations, which frequently collect more data than they use in decision-making. Reasons for this practice include, uncertainty about what sort of data will be relevant to decision making, the

desire to have data that can be used to justify a decision after the fact, and the desire to display a symbolic commitment to evidence-based decision making.

It is even more difficult to determine when action has been prompted by the process of producing or reacting to an indicator. The relationship between cause and effect is not always clear – does the process of producing an indicator cause mobilization, or is it the reverse? In the case of state failure indicators, causality has worked in both directions: a CIA-sponsored effort—the State Failure Task Force (1994-2000)—to identify the determinants of state failure and instability did not only produce data and analytical techniques, for the CIA and other users, but also 'mobilized' a group of scholars who have gone on to become involved in the production of other indicators of state failure.

5.2. What factors determine the influence of indicators?

In the absence of hard evidence on the influence of indicators we can make informed guesses about factors that will determine their influence. Both the direct influence of indicators as forms of information, and the influence they exercise indirectly through social mobilization have to be considered.

Information channel: demands for indicators as information

Indicators can be thought of not just as forms of information, but as *informational products*, like news, or weather forecasts. The question of what makes indicators influential among potential actors then becomes: what factors shape the demand for indicators? Conventional economic theory suggests some answers.

The economic analogy suggests that a key determinant of the influence of indicators will be the cost of accessing them. Indicators that are assimilated in the course of acquiring information for general purposes and succeed in becoming part of potential actors' background knowledge are essentially costless to access. Indicators that are published on prominent freely accessible websites, that are easily found using standard search techniques, and that are presented in a user-friendly format, are also relatively inexpensive to access.

The other parts of the equation for calculating the demand for, and hence influence of, indicators, are the benefits people expect to flow from going to the effort of accessing them. Potential actors should, all other things being equal, gravitate to sources that point toward actions that generate better outcomes.

However, the intrinsic features of an indicator only partially explain the perceived benefits of relying on it. As with other products, much of the perceived value, and thus demand, will also depend on the context in which the indicator is introduced. The demand for a product typically depends on the availability of related products and the number of people who are already familiar with it. Similarly, the demand for any given indicator will depend on the

availability of other potential sources of information and the number of people familiar with it.

One particularly important factor is the availability of *substitutes* for the indicator in question. The demand for any given indicator for any given purpose will depend on the availability from other sources of information that can be used for similar purposes. All other things being equal, the demand for an indicator will tend to fall as its substitutes become more accessible.

The idea that demand for indicators is negatively related to the availability of close substitutes implies that *product differentiation* will tend to increase demand for an indicator. Prospective users are more likely to demand an indicator if it provides 'surprising' information, that is to say, information different from that generated by other sources of information, whether those sources are other (cheaper) indicators, forms of information that are not indicators, or the users' pre-existing knowledge. For example, the producers of the Failed State Index report that their indicator attracts attention when it ranks Chad close to well-known failed states such as Somalia because of the "shock value" for readers unfamiliar with the deteriorating political situation in Chad. There is, however, a fine line between an indicator that is sufficiently novel to be informative and one that is so out of step with existing knowledge as to be incredible.

Demand for indicators will tend to rise as their substitutes become less accessible. In fact, the absence of other sources of information can lead to indicators being used for purposes quite different from the ones originally intended by their promulgators. For instance, the WHO/UNICEF immunization coverage indicators are sometimes used as measures of the overall quality of national health care systems. This situation may be explained in part by the absence of any other single indicator of national health system performance suitable for making cross-country comparisons. In 2000, WHO published *The World Health Report 2000 – Health systems: Improving performance*, wherein it measured health systems in 191 member states on the basis of five indicators: overall level of population health; health inequalities (or disparities) within the population; overall level of health system responsiveness (a combination of patient satisfaction and how well the system acts); distribution of responsiveness within the population (how well people of varying economic status find that they are served by the health system); and the distribution of the health system's financial burden within the population (who pays the costs). The political uproar that followed (prompted not the least by the U.S. ranking 37th) caused WHO to stop any subsequent explicit ranking of health systems. Immunization coverage indicators have filled the gap.

The demand for an indicator may also depend on the availability of products that are *complements*, meaning, products that increase the benefits to be earned by relying on the indicator. Among those complements are tools for analyzing the data contained in indicators. This may be why the producers of GIIRS also produce software that allows users to see the effects on their investment portfolio's social impact rating of adding a company or fund to the portfolio and to see how a particular company or fund performs in comparison to its peers.

In some cases the value of an indicator will depend upon how popular the indicator is. This will be true whenever the use of an indicator generates positive *network effects*. The classic example of a product whose use generates network effects is a telephone. The more people who already have telephones, the more valuable the telephone is to other users; each additional owner of a telephone makes the telephones of other owners that much more valuable. A typical feature of indicators is that the more actors who report their scores using a particular indicator, the more the indicator can be used for comparative purposes and therefore, the more valuable the indicator. So for example, the more investment funds there are that report the social impact rating of their portfolios using a particular indicator, the more funds that can be compared to one another. Thus the success of GIIRS – or any other social impact investment indicator for that matter – is likely to turn on whether it is used for reporting purposes by a critical mass of investment funds.

The demand for an indicator at any given point in time may also depend upon how popular the indicator is anticipated to be in the future. This will be the case when use of the indicator generates positive *learning effects*. Products generate learning effects when their use provides benefits for future users. A good example is the QWERTY keyboard. The more people who have used QWERTY keyboards in the past, the easier it will be to find people have learned to type quickly on it and so the greater the value of the keyboard to a firm that purchases it. Similarly, the larger the supply of people who understand how to use a particular indicator in decision-making, the easier it will be for an organization that relies on the indicator to find staff. This may explain why the producers of the GIIRS social impact investment indicators have made a concerted effort to induce major investors to pioneer the use of their indicators.

Social mobilization channel

Indicators can be thought of as processes for mobilizing actors as well as sources of information. The indirect influence of the indicator will presumably depend on the number of actors who are involved in the production process, and the power those actors have to influence others, whether by virtue of their centrality in various kinds of social networks, their economic power, or their perceived expertise. This perspective suggests another explanation for why GIIRS has strived to involve so many influential actors in the process of developing its social impact investment indicators.

5.3. Under what conditions can indicators promote optimally constructive action?

Action through information

Accurate indicators should, to the extent they are influential at all, help actors identify the most constructive ways of addressing problems. The reverse is also true: inaccurate indicators point toward less constructive actions. Accurate social impact investment indicators will encourage socially-motivated investors to allocate funds to firms whose operations have the greatest social impact. Inaccurate indicators, however, will channel

funds toward firms that have less positive impact than firms that are neglected. Similarly, inaccurate indicators that overstate the extent to which societies suffer from vaccine-preventable diseases wrongly point donors and national officials towards prioritizing more immunizations over other health system interventions.

There is no guarantee, however, that actors who use indicators will use them to address social problems. For instance, to the extent that they provide information about the future many indicators can be used to profit from rather than to change the course of events. Thus, indicators that predict a famine can be used to speculate in commodity markets as well as to plan relief efforts. Similarly, investors can use indicators of state fragility to decide which countries to avoid—avoiding loss is equivalent to increasing profits—rather than to identify regimes that might benefit from constructive engagement.

Indicators can also be used self-interestedly to control the behavior of the actors whose behavior is being measured. For example, fiscal indicators such as the level of a country's budget deficit might in principle be used to measure and address problems associated with waste of government resources. But those indicators can also be used by a country's creditors as part of their efforts to control the government's spending in order to ensure that their loans are repaid.

Action through mobilization

Indicators can influence action by mobilizing actors around either constructing or contesting them. Whether the actions that result from this kind of mobilization are constructive depends on the motivations of the group members and their collective knowledge and expertise, including the kind of expertise that comes from direct experience with the types of disadvantage the group aims to address.

A case in point is the network that surrounds the GIIRS social impact investment indicators. GIIRS is produced by B Lab, a U.S. based non-profit. However, GIIRS relies heavily on a data collection template created by IRIS, an initiative of B-Lab, Acumen Fund (another non-profit) and the Rockefeller Foundation. IRIS is now a project of the Global Impact Investing Network (GIIN), a nonprofit organization (incubated by the Rockefeller Foundation and launched at the Clinton Global Initiative in September 2009) that is dedicated to increasing the effectiveness of impact investing. The main governing body of the GIIN, the Investment Council, brings together about 30 institutions involved in the impact investing industry and made up of private foundations, major banks, institutional asset managers, family offices, and pioneering impact investors. In the development of GIIRS B-Labs relies heavily on Standards Advisory Councils (SAC) made up of experts from different areas, including non-profits, academics, performance evaluators, social finance, international development, and so on. The SACs convene quarterly as a group and, at least during the development period, members of the GIIRS staff have been in constant communication with individual members. In the future GIIRS envisions developing partnerships with other impact investing organizations such as the More for Mission Campaign, the Singapore Stock Exchange, as well as with conventional rating agencies. Whether or not the creation of this network results in significant expansion of interest in social impact investment depends on whether the

members are committed to the effort or are simply attempting to burnish their image, as well as whether they collectively possess the expertise to identify the best ways of promoting social impact investment. It remains to be seen whether the technical expertise and financial resources available to members of the group can compensate for the relatively low level of involvement on the part of potential beneficiaries of social impact investment.

5.4. When do indicators promote coordinated action?

To the extent that indicators foster common understandings of a problem they are likely to make the decisions of actors who rely on indicators more *consistent*. As we have already discussed, however, it is not clear that producing indicators necessarily fosters common understandings. Moreover, *consistency in decision-making is not an unalloyed good*. Imagine for example, if the bulk of major social impact investors use the same indicator to determine which firms are eligible to receive funds. Or alternatively, suppose that collaborating to produce such an indicator leads the investors to share beliefs about what kinds of social impact are most worth funding. Either way, the result is likely to be a stampede to provide capital to the handful of firms at the top of the rankings, leaving firms that are only slightly less eligible than the investors' darlings out in the cold.

Uncoordinated reliance on indicators may not be a serious problem if indicators can be updated frequently to take into account the effects of recent events. So for example, the danger of investors flooding highly rated firms with more capital than they can use diminishes if social impact investment ratings are updated frequently to reflect how inflows of capital change firms' projected performance. Sometimes, however, changes in the accuracy of indicators take a long time to observe. Moreover, frequent updating of indicators can be costly.

5.5. Does use of indicators enhance accountability for actions?

For some actors—typically large organizations with external stakeholders, such as government agencies or public charities—it is important to be accountable for their actions, including actions aimed at addressing social problems. An important determinant of accountability is transparency, and indicators have the apparent virtue of permitting relatively transparent decision-making—at least at first glance.

The transparency associated with indicators stems from the fact that they are easier to communicate than other less heavily filtered kinds of data—it is easier to communicate a country's score on a failed states indicator than it is to transmit the in-depth multi-faceted analysis of its political climate on which the indicator is based. Ease of communication implies that it is easier for people to access the direct informational basis for decisions based on indicators than it is to access decisions based on less filtered sources of information. This is the sense in which making decisions based on indicators tends to be transparent, at least in comparison to those based on raw data.

The transparency of decision-making based on indicators is merely superficial, however, if the process of producing the indicator is opaque. In some cases it is not even clear what data are incorporated into the indicator. This is most evident when the data consist of the opinions of anonymous analysts, whose knowledge and training are impossible to ascertain. In other cases, there are so many different and obscure sources of data that it is impossible to evaluate them. For example, the Institute for State Effectiveness, a Washington-based think tank, has proposed a sovereignty index based on 10 "core functions" that states should perform, which will be measured using 100 indicators. In this case reliance on the indicator may make it more difficult to identify and obtain the data that provides the underlying foundation for a decision.

In still other cases, the fundamental problem is failure to disclose the methods used to aggregate different sources of data. The Failed States Index is a classic example. Scores on each of the 12 sub-indicators used to calculate the overall index are calculated using an undisclosed computerized method of analyzing various publicly available sources, followed by review and adjustment by in-house analysts.

6. Contestation around indicators

Indicators that qualify as interventions risk attracting contestation. This takes place at many levels. Where different units (such as states) are ranked, the leaders of units that rank poorly will often protest the methodology or data used, or the choice of criteria on which the evaluation is made, or (where the entity did not specifically consent to be ranked) the presumptuousness of the promulgator in arrogating to itself the power to issue such rankings. More broadly, if the indicator seems likely to have significance for what is valued or for future policies and resource allocation, people who do not share the framing, the causal theories, or the set of desired actions embodied in the indicator may protest it. People may protest for other reasons having to do with the propriety of the promulgator exercising this power, or with the process by which the indicator was produced (matters such as transparency, consultation, participation, reason-giving, review; or issues concerning how and by whom the data was collected, or what was excluded). Contestation may extend beyond protest, to other strategies such as production of competing indicators, or efforts to discredit the indicator or its promulgator, or promotion of alternative framings and alternative priorities for action.

The World Bank's Doing Business indicators have provided a basis for both specific and more general contestation. One particular component of the Doing Business set of indicators, namely, the Employing Workers Indicator (EWI) has been particularly provocative. The EWI was subject to sustained challenge from organized labor groups, working in part through the International Labour Organization (ILO). They objected that the EWI seemed to reward countries for making it easy to fire workers, and that the policies which attracted high EWI rankings were often radically at odds with standards embodied in ILO Conventions. These opponents received support from key members of the US Congress. In the Doing Business Report for 2011, the International Finance Corporation (the entity in the World Bank group which produces the indicators) ceased to give any weight to the EWI, while preparations were made for a replacement indicator much more closely aligned to ILO Convention standards.

The Doing Business indicators have also been contested for more general reasons. Many Bank personnel and outside commentators have expressed private or public skepticism about relying on the indicators to make policy. It is reported that at the October 2010 meeting of the World Bank's Executive Board, executive directors representing Brazil and China were among those expressing opposition to the Bank's continued use of rankings in this area. Rogerio Studart, the executive director representing Brazil and several other countries on the Bank's board, criticized the Doing Business report on methodological grounds, but also on the basis that the report reflects a particular ideology and privileges one regulatory approach above others without sufficient justification. He is reported to have said: "I've always been struck by the exuberance of the propaganda they made out of it and the pressure they would put on some governments by using the rankings to adopt reforms, as if those reforms would solve some fundamental problems that in my view they could not

solve. Reducing the number of procedures and the number of days to open up a company -- this is always helpful. But portraying that as a way to create more jobs is a total jump." (Bosco 2010)

The opposition to the Doing Business indicators exemplifies the intense suspicion some global governance indicators have inspired within developing countries. Such indicators reflect theories about the proper governance of societies that are difficult to challenge because they are not stated openly, and are only evident on close analysis of the measurement and aggregation processes that go into production of the indicator. As developing countries often lack the expertise and resources to engage in this close analysis, use of such indicators is resisted precisely because of concern that it marginalizes dissenting arguments from those countries in policy debates about how best to improve outcomes for their populations. Concern about these effects is intensified because many global indicators are poorly tailored to the circumstances of developing countries, and can exercise particularly pernicious influences in relation to fragile and conflict-affected states. Objections to global indicators are particularly strong when they are suspected of having a distorting effect on local policy decisions. This may be because performance on certain indicators is a precondition to accessing external aid, or because promulgation of the indicator aims to give added weight to demands for particular policies by small elite political constituencies within developing countries or by locally-connected external actors such as prospective foreign investors.

The case of the Doing Business indicators demonstrates that production of an indicator can influence both the balance of power in contests to shape knowledge about social problems and possible strategies of contestation. Production of an indicator that supports one set of beliefs typically places the burden on proponents of alternative beliefs to produce additional information. Whether or not they can meet that burden will depend on many factors, including features of the indicator being contested. For one, producers of the indicator make it costly for their rivals to obtain the raw data required to become a credible source of information about the topic. This will in turn depend on whether the producer of the indicator has legal or de facto control over the underlying data and how it exercises that control. Indicators that crowd out alternative sources of information will also make contestation difficult.

This all suggests that the distribution of power to shape knowledge will be influenced by who is best placed to produce and contest knowledge based on indicators. Some actors are better placed to shape knowledge through indicators than others. In general these will be actors with superior abilities to collect, process, distribute and publicize the existence of data. These are typically organizations with financial resources or social connections that give them privileged access to certain sources of data, ready access to social scientific expertise, and high public profiles. It is less clear however, which actors are best placed to contest knowledge based on indicators. To the extent that it takes an indicator to beat an indicator then the ability to produce indicators will shrink the group of contestants to those who are capable of producing indicators. Further research is required, however, on which sources of information are best placed to compete with indicators.

7. Learning and revision

Addressing complex development problems in many cases depends heavily on local action and innovation. While some basic principles and parameters may already be clearly established and broadly applicable across all sites, many essential elements for success may be quite uncertain, or indeed unknown. It may therefore be very desirable to promote locally-initiated experimentation at numerous sites. Externally-supported development intervention can focus on making such experimentation possible, structuring processes of comparison, learning, and dissemination of insights from more and less successful innovations and approaches, and supporting concerted action in relation to external political institutions and donors. Local actors can then adapt or reconstruct their approaches in light of these insights to improve their own operations and outcomes. This process of learning and revision and can continue over a sustained period as new ideas emerge, and also to build resilient approaches as experience arises with sudden local adversities or with large-scale shocks and changes of conditions affecting multiple sites.

Indicators can be a key component of such initiatives to improve practices on complex issues through experimentation, learning and revision. Data supplied in comparable form from multiple sites, with any necessary monitoring and verification, can be compiled into indicators and distributed back to the local actors. These actors should also have opportunities to share experiences and exchange views through narrative reports, videos, discussion forums, meetings etc. Indicators are used extensively in such experimentalist learning within the European Union, particularly under the Open Method of Coordination and comparable modes of governance.[4] These uses engage a wide range of participants, including groups of directly affected persons, labor unions, employers, local and national government agencies, and transnational networks of such organizations. This structure of governance on issues of social inclusion, can thus be in itself a means not merely to measure inclusion or indirectly to spur action, but also to help directly in increasing inclusion and agency. Some such initiatives have been pursued also within the OECD, and within particular developing countries.

It does not seem that global governance indicators involving North and South have been used very much in this way. Some important examples exist, including efforts to promote climate change resilience in developing country cities that seek to evaluate experiences in ways that facilitate learning and transfer of promising approaches between sites where comparable conditions appear to exist. Further research is needed on the roles of indicators in such areas of practice. Investigation is also required as to the degree to which it has been (and may in the future be) significant that indicators produced by global organizations on development issues function in an environment characterized by power and wealth disparities and a degree of mutual North-South mistrust. It is possible that this leads indicators to be used in governance contexts of performance evaluation or advocacy rather

[4] Charles Sabel and Jonathan Zeitlin, eds., *Experimentalist Governance in the European Union: Towards a New Architecture* (Oxford: Oxford University Press, 2010).

than experimentalist learning. Even if this is correct, further active piloting of such indicator-based experimentalist governance within and across developing countries appears a promising direction for future initiatives.

Framing, action, and contestation provide experience and ideas which become learning. Learning leads to the need for updating: of basic framing of issues, of ideas and practices concerning actions that prove positive or negative, and of practices of contestation and objectives sought in contestatory processes. The data incorporated into indicators itself provides one basis for learning, although much more than numerical data is usually needed. A more complex question is whether the indicators themselves ought to be revised, to reflect new framings, new ideas about causes of problems and about pathways of action, and learning from contestation of the indicator or of theories embodied within it. Major revisions of an indicator, including major changes in what is measured and how, may impair comparisons across time, and they may also be costly to implement accurately.

While major global indicators are frequently subject to minor adjustments, we have not found examples in case studies of immunization coverage, state effectiveness, and impact investing indicators of fundamental revision resulting from new thinking precipitated by a learning process. One reason for this is that, two of the three indicators in our case studies have been created relatively recently, and there have not been fundamental shifts in thinking since any of the three indicators were established in their current forms. This applies particularly to impact investing, for which the indicators are still under development. There are signs of specifically-focused change with regard to the issues concerning failed or fragile states, precipitated by contestation and deepening knowledge, which may quite rapidly come to be reflected in the production and use of indicators. Some revision in broad approaches to health metrics may eventually have implications for immunization coverage indicators. The levels of contestation and rethinking around the existing immunization coverage indicators have not been high, but the demand for more fine-grained context-specific measurements and interventions is strong, as is pressure for greater emphasis on wider health systems approaches.

8. Conclusion

Indicators are used in addressing development problems ranging from corruption to human trafficking to obstacles to doing business to state failure. Indicators are routinely used to monitor and evaluate the impact of various kinds of development interventions – donors often require them. Indicators are also used at times simply for their advocacy effects. What is sometimes overlooked is that many such indicators are not simply tools for diagnosis or evaluation or advocacy – they are also interventions in their own right. Creating an indicator can be a constructive intervention (or in some cases a counterproductive response) in addressing a 'wicked' social problem. This report has sought to show the potential of indicators as constructive interventions in efforts to address complex development problems; and to draw attention to some pitfalls and perverse effects that can easily result from producing and using indicators without realizing that they are themselves interventions, or without sufficiently far-reaching thought and consultation.

Indicators can be significant interventions in addressing development problems because they can change beliefs, including beliefs about the moral implications of various social outcomes, relationships between actions and outcomes, and, the beliefs or actions of other actors. In doing so, indicators can affect the decision-making process, actions and, eventually, outcomes.

Well-crafted indicators can be particularly effective in addressing wicked problems of development by creating shared understanding of the problem among diverse actors. People who are exposed to the indicator may come to accept the claims about problem framing and standards that the indicator embodies. In addition, both the process of construction and dissemination may change the views of at least some of the participants. The case studies in this report suggest that common understandings have been promoted by the WHO/UNICEF's processes in production of immunization coverage indicators and the substantial groundwork done by the promulgators of GIIRS may produce some shared understandings among certain categories of investors, certainly in North America and perhaps also in the North Atlantic more broadly.

Indicators can do more than just influence understandings of problems, they can also influence action by alerting people to the existence of a problem, helping them to understand its magnitude, and pointing them toward means of addressing it. Our case studies and literature review provide examples of indicators that have influenced action by intergovernmental and governmental leaders and agencies, donors and voters, public officials, for-profit corporations, volunteers and not-for-profit organizations, alliances, or networks; and, of central importance, members of affected communities. WHO/UNICEF immunization coverage indicators are perhaps the most prominent example, with donors using the indicators to decide how to allocate funding for global health initiatives. In other cases, the causal influence of indicators is not as clear. For instance, the creators of the Failed State Index claim that it has been referred to in foreign aid allocations, but it is difficult to determine the extent to which data from this indicator guides ultimate actions. This is true

of other indicators as well. Actors often rely on many different kinds of data in the course of deciding whether and how to address problems, and they do not necessarily understand, record or disclose the elements of their decision-making processes.

Some of the factors that might determine the influence of indicators include the cost of accessing them, the benefits people expect to flow from going to the effort of accessing them, the context in which the indicator is introduced (such as the availability of related products and the number of people who are already familiar with the indicator), and the availability of other familiar potential sources of information that could be used as complements or substitutes for indicators. Demand for indicators will tend to rise as their substitutes become less accessible. In fact, the absence of other sources of information can lead to indicators being used for purposes quite different from the ones originally intended by their promulgators, as the case study of WHO immunization coverage indicators demonstrated.

In some cases the value of an indicator will depend upon how popular the indicator is. This will be true whenever the use of an indicator generates positive *network effects*. The success of GIIRS, for example, may turn on whether it is used for reporting purposes by a critical mass of investment funds. The demand for an indicator at any given point in time may also depend upon how popular the indicator is anticipated to be in the future. Thus, making a concerted effort to induce major investors to pioneer the use of their indicators is likely to contribute to the success of the GIIRS social impact investment indicators.

Indicators can encourage a network of actors to address the problem constructively through the social and political engagement they foster, independently of their direct effects on people's beliefs. Producers and advocates of indicators can form networks, as can actors who challenge or criticize indicators. Both sets of networks can span multiple societies, enabling broad dissemination of information, which in turn results in opportunities for new understandings of the problems, appreciation for causal relationships between actions and outcomes, and exchange of information about past or anticipated actions. These networks can also foster cooperation across a range of actions.

Finally, indicators can be a key component of initiatives to improve practices on complex issues through experimentation, learning and revision. Addressing complex development problems in many cases depends heavily on local action and innovation. The data contained in indicators can record the effects of locally-initiated experimentation and play an important role in processes of comparison, learning, and dissemination of insights from more and less successful innovations and approaches. Local actors can then adapt or reconstruct their approaches in light of these insights to improve their own operations and outcomes.

Appendix

Producing and Using Indicators as Interventions in Complex Social Problems: A Checklist

1. Framing the social problem

   WHO?

   _ Who should be defining the problem?
   _ Who should be participating in indicator design and use?
   _ What consultative processes are required?

   WHAT?

   _ What definitions of the complex social problem are already framed?
   _ What are some alternative ways of framing the problem?
   _ What criteria should be used in choosing the framing to use?
   _ What existing or potential new indicators align with this framing? Are they narrower, overlapping, or broader than the problem as framed?
   _ What is the likely contribution of using a particular indicator?
   _ What might detract from the indicator's utility, or pose a risk of negative consequences from the indicator's use?

   WHY?

   _ Why produce an indicator (what will it add to existing interventions?)
   _ How will the indicator interact with existing activities?
   _ Will the indicator support or undermine other possible interventions (e.g., legislative advocacy, court challenges, video production, narrative reports, etc.)?
   _ What are the perspectives of the intended beneficiaries of the intervention? Of those whose activities and performance will be measured by the indicator?

2. Designing the indicator
   _ Who are the intended beneficiaries of the indicators?
   _ What are the units to be measured?
   _ Who are the relevant audience for the indicators?
   _ What are the intended uses and impacts of indicators?
   _ What are the possible unintended uses of the indicator?
   _ Who will be impacted by the indicator?
   _ How will this framing and use of the indicator be perceived by the units or persons being measured? How will it affect these units or persons and their behavior and attitudes?
   _ How will this framing and the possible uses of the indicator be perceived by, and affect, the intended beneficiaries of the intervention?
   _ What are the likely or possible effects on factors affecting the social problem as a whole?

3. Producing the Indicator

- _ What are all the possible sources of data? What are the strengths and weaknesses of data currently available or likely to be available from each source (e.g., cost, reliability, completeness, etc.)
- _ How will missing data be deal with? What are the downsides of this?
- _ Can adequate safeguards be instituted to deal with intentionally misleading/inaccurate data?
- _ What methodologies are to be considered for construction of the indicator? Which methodology is likely to yield the most accurate, valid, and useable indicator? How will evaluation of this choice be conducted?
- _ Which stakeholders ought to be included in indicator production?
- _ Is the methodology transparent, verifiable and replicable by independent parties? In particular, can the methodology be understood and replicated by (a) units measured and (b) intended beneficiaries of the intervention?
- _ How might the process of indicator production affect the utility of the indicator? (e.g., opaque production process can thwart the aim of making the indicators the basis for decision-making)
- _ What degree of uncertainty is attached to the indicator?
- _ Are safeguards needed to curb gaming of the indicator, or to ensure that it does not undercut the overall approach to improving the complex social problem?

4. Using the Indicator

- _ Who is using the indicator (and who is not?)?
- _ What information is conveyed to the different users by the indicator and related material and debates?
- _ Is this information having effects on users' beliefs about the world?
- _ Is it crowding out other ideas and data sources?
- _ Are networks forming around the production, use or contestation of the indicator?
- _ Are these networks contributing to addressing the social problem, and in what ways?
- _ Are there any counterproductive effects of the indicator on social mobilization (e.g. promoting rivalry that undermines cooperation; undervaluing or alienating key actors)?

5. Enabling contestation and learning once the indicator is in use

- _ Is a process in place to review periodically not only progress on the complex social problem, but also to review the value, uses, and effects of the indicator as an intervention?
- _ What processes are in place to ensure insights of different people and groups affected are articulated and heard?

6. Learning from experience in production and use of the indicator

   - Have the indicator or the framing of the social problem associated with it been subject to direct contestation? Are concerns being expressed indirectly about it, and how can these be translated and engaged with?
   - Is the indicator in fact measuring what it was intended to measure?
   - Have there been unintended (positive or negative) impacts of the indicators?
   - Should the framing of the problem as reflected in the indicator be re-evaluated?
   - Is revision of the indicator necessary, for example to reflect learning or new events? (Taking into account costs involved in revisions, and effects that frequent revisions might have on the utility and reliability of the indicator.)