Evaluating Social Change and Communication For Social Change: New Perspectives

By Ailish Byrne

Today's growing interest in assessing and demonstrating the impact of social change and communication for social change processes is the focus of this essay by Ailish Bryne, the Consortium's senior associate of research and evaluation. Using the meeting of HIV/AIDS Implementers in Kampala, Uganda, last June, Byrne lays out the challenges and diverse factors impacting measurement and evaluation. She raises questions regarding many current M&E practices and their implications for development.

> The received wisdoms of M&E are being fundamentally challenged based on a different understanding of development itself as complex, emergent and transformative. However these shifts in thinking are slow and much is needed to come to assessment and learning processes that strengthen social change rather than hinder it. (Guijt, 2007: 51)

> Human systems are dynamic, entangled, scale independent, transformative and emergent. These characteristics challenge the basic assumptions of traditional evaluation methods. They necessitate new evaluation approaches that are as rich and varied as the human systems they are designed to assess (Eoyang & Berkas: 9)

> A new paradigm has transformed the evaluation environment and raised the bar for the fledgling development evaluation profession. In the new century, development evaluation will have to be transformed and liberated from its current strictures to reach beyond aid, beyond projects, and beyond top-down approaches. It should be more comprehensive, more participatory, and better adapted to the felt needs of society. (Picciotto: 520)

Traditional M&E needs complementing with other approaches, particularly in relation to social change and communication for social change. This does not imply more M&E but, rather, doing what we do *differently and better*. This essay

highlights issues that continue to impede, rather than foster, progress. I attempt to show the need for deeper reflection and questioning on the part of senior development stakeholders in particular. And I call for more openness to methodologies that seem to me more appropriate to evaluation of social change.

I. The Kampala Conference and Social Change Communication
In June 2008 I attended the HIV/AIDS Implementers Meeting in Kampala on behalf of the Consortium. The meeting was hosted by the Government of Uganda and co-sponsored by the U.S President's Emergency Plan for AIDS Relief (PEPFAR), the Global Fund to fight AIDS, TB and Malaria, UNAIDS, UNICEF, the World Bank, WHO and the Global Network of People Living with HIV/AIDS (GNP+).

I was a panel member presenting on *Communication for Social Change: Community Dialogue and Partnerships Impacting Social Norms about AIDS*. The theme of the Implementers' Meeting was "Scaling Up Through Partnerships: Overcoming Obstacles to Implementation". More than 1,700 participants from 70 countries attended.

In his opening speech Mark Dybul, PEPFAR Coordinator, noted that "seemingly ordinary people who appear to be without hope can do extraordinary things with a little support".

Dr Michel Kazatchkine, of the Global Fund, emphasised the importance of sharing lessons to control the pandemic better. In the context of the meeting's goals, this raises questions about why evaluations are so rarely designed to capture and legitimise the efforts and achievements of "ordinary people" and why evaluation practice conducive to sharing learning is not widely advocated, particularly given the notable potential of evaluation.

The social change communication (SCC) panel included leading practitioners in this field, united by their belief in the power of communication to address the social drivers of HIV/AIDS, including socio-cultural norms that impact an individual's behaviour and degree of choice.

We defined SCC for AIDS as "the strategic use of advocacy, communication and social mobilisation to systematically facilitate and accelerate change in the underlying drivers of HIV risk, vulnerability and impact". Our focus was HIV/AIDS programmes and prevention strategies, where the limitations of widespread individual behavioural change approaches, which pay insufficient attention to social and cultural norms and contexts, continues to present a major communication challenge.

The session considered: What can be done to catalyse shifts in harmful social norms? How are sustainable social change strategies implemented? What can we learn from experience to date? What can be learned about SCC in other

fields that can be applied to AIDS? What are the key challenges? What can be done in the immediate, short- and longer-terms?

We emphasised the importance of developing and sustaining alliances and partnerships, of sustained capacity development and support, of sharing results and lessons learned, to progress the field. The importance of strengthening and actively supporting the M&E of social change communication was emphasised.

The session was grounded in extensive panel experience and rich discussions over time amongst members of the UNAIDS SCC Working Group. We attempted to model a SCC process in the session itself, as far as possible; to give participants a "taste" of SCC approaches that can change destructive social norms.

Although ambitious and, inevitably, limited by time and the artificial setting of the conference, the relative interactive format was appreciated by participants. I am grateful to other panel members for the constructive ideas and discussions shared, however these reflections are entirely mine

An untimely death framed and gave the session added depth and poignancy. Wellington Solomon Adderly, a prominent AIDS activist in the Bahamas, was killed at home days before the Kampala meeting. This tragedy powerfully represented why our work in SCC remains so important and urgent: How could a SCC approach help ensure that such a tragedy never happens again?

In groups, participants discussed real-life HIV/AIDS scenarios and determined the implications for communication initiatives necessary, in short- and long terms. The vignettes had been chosen to highlight issues of vulnerability, poverty, gender, violence, concurrent partnerships, lack of choice, stigma and discrimination, amongst others.

Participants engaged enthusiastically and many later commented positively on the depth and breadth of panel members' contributions and experience. In particular, people appreciated framing the issues within multiple timeframes (immediate to longer-term), emphasis on multi-sectoral and multi-disciplinary perspectives and partnerships, on joint responsibility for the solution, on the importance of comprehensive approaches, on community-level interventions and framing it in a rights-based approach. Many requested more case studies of SCC and its evaluation. And some said SCC should have been covered in plenary.

The process of developing the session, as well as feedback afterwards, reaffirmed the need for more concrete examples of the impact of SCC. Sound evaluation and demonstration of impact is urgent; in particular to convince key stakeholders, including donors, of the value of SCC approaches. Unpacking this oft-repeated plea raises, for me, fundamental questions about why effectively demonstrating the impact of SCC appears so challenging and difficult to achieve

in practice. In turn, this begs questions about factors that continue to impede progress in this area and those that would strengthen it.

Questions include why it remains so difficult to negotiate and fund innovative evaluation practice, why there appears such resistance to changing dominant M&E practice at senior levels and who needs to change.

The implications for organisational cultures, dominant research paradigms and practice and "expert" committees are considered: Who makes influential decisions and on what basis? What constitutes appropriate data? Who defines this and whose interests are ultimately served? Underlying the discussion is a plea to broaden and *complement* (rather than replace) current M&E practice.

These critical issues are returned to below, following attention to the characteristics of social change that make its evaluation inherently complex.

This article attempts to respond to deeper questions underlying my article in *Mazi* November 2007 as well as in Virginia Lacayo's article in *Mazi* August 2008.

At the risk of over-generalisation, certain characteristics remain true of most M&E practice in international development and communication.

- How have we come to a situation in which many people and organisations in the field feel continually frustrated at having to fit their achievements and efforts into externally-defined and imposed "SMART" objectives and indicator tables like Logical Frameworks?

- Why does there appear so much "evaluation for evaluation's sake", rather than evaluation for improvement and learning?

- How do certain institutions and individuals come to be seen and legitimised (by powerful stakeholders) as "experts" or authorities in the M&E field, to the neglect of other possible experts?

- And why is the authority of traditional "experts", so readily assumed and rarely questioned, even when experience repeatedly demonstrates the limitations and weaknesses of what they advocate?

 I am continually reminded that "you find what you're looking for"; if you search for more of the same, you will get more of the same. We have to ask why powerful stakeholders appear so resistant to broadening their horizons and learning from innovation and experimentation elsewhere.

The point of evaluation is generally to capture as effectively as possible the impacts, both intended and unanticipated, of particular initiatives. This can be done in ways more or less conducive to reflection, analysis and learning, at

individual and organizational or collective levels. It does not make sense to insist on "perfect measures" whose appropriateness and feasibility will long be contested, particularly where social change is the focus.

Debates rage about the possibility and indeed desirability of objectivity in evaluations, because such notions clearly favour one perspective and actively negate others. The implications are profound and the questions raised uncomfortable. But they are ones that must be asked by those of us who seek to address imbalances in voice, opportunity, resources and support, at local and international levels. If one hundredth of the effort and expense devoted to M&E over the years had been spent on alternative, innovative approaches of the kind advocated below, we would already have much more, sound evidence of the impact of social development and communication initiatives.

So why is this so hard to come by?

It is evident in the major effort currently under way to channel scarce funds towards experimental, *i.e.*, traditional evaluation methods and push these towards the South. This follows a high-profile U.S. government campaign that "enthroned RCTs [randomised control trials] as the gold standard for evaluation" (Picciotto). Controversial as this scheme is, it illustrates how powerful and non-neutral forces actively perpetuate and impose the status quo, often for dubious reasons, while actively inhibiting innovation and healthy experimentation in the process.

Alarmed at the implications of such trends, Picciotto recalls the major debate amongst evaluation scholars two decades ago when the mainstream evaluation community carefully delineated conditions under which experimental and quasi-experimental methods are appropriate: "They are useful only for discrete and relatively simple interventions geared to precise and measurable objectives and characterised by well-defined 'treatments' that remain constant throughout program implementation" (Picciotto: 517). Clearly, these characteristics do not apply to the complex, multi-faceted and largely unpredictable processes of social development, social change or SCC, so their value to the evaluation of such initiatives is at best limited.

Prowse (2007) similarly outlines the limitations of RCTs in the context of multi-faceted, multiple development processes, highlighting their questionable assumptions of objectivity and mirroring the natural world and their neglect of human agency, interpretation and complexity. His wider study of high-level literature on Impact Evaluation shows alarming bias towards Randomised Control Trials (RCTs) and neglect of qualitative methods.

Policy makers, donors and large aid organisations are key actors in this state of affairs and there are growing calls for critical assessment of the evaluation practice they advocate. As recent research into global efforts to assess social

change highlights, "donors, by and large, favour a mode of M&E rooted in fears of non-compliance of agreements based on a development model that is considered predictable. In practical terms, donors need to rethink the principles on which they base their models of evaluation and learning. Amidst what might seem like a daunting agenda, one action point merits special attention, that of consistency—donors must be more rigorous in aligning their espoused values with the protocols and systems they use" (Guijt, ASC: 8).

This is not to suggest that donors are not themselves subject to accountability to higher authorities (Eyben: 14), are not also frustrated by much current practice, or that all are averse to considering and funding alternatives (there are notable exceptions). A recent international survey of aid donors and recipients shows that both groups prioritise strengthening impact assessment. However, "donors are not funding it adequately or making effective use of results, nor are they investing in a larger infrastructure to support it" (Bonbright, 2007: 2).

The evaluation of SCC remains notably weak for many reasons, including lack of investment and challenges raised by the essence of social change itself. However my arguments have wider relevance.

Shifts in the HIV/AIDS field: Towards a social perspective
Returning to the theme of the Implementers' Meeting, HIV/AIDS prevention serves to illustrate why time is ripe for SCC approaches, urgently and at greater scale. This will not happen without better demonstration of impact.

The costs of failing to appreciate the significance of socio-cultural factors and norms like multiple concurrent partnerships, circumcision, gender violence and inequities to HIV prevention, are increasingly realised. This was a theme in Kampala and was highlighted at the Mexico XVII International AIDS Conference in August 2008.

Warren Feek of the Communication Initiative has referred to Mexico as the "social conference" because of the focus on social drivers, social complexity, social change, social mobilization, social movements, social stigma, socio-economies, socio-cultural factors and social phenomenon. This is reflected in *The Lancet's* August 2008 issue on HIV prevention, released in Mexico. It includes data and conclusions stressing the need for a long-term approach, for programmes grounded in and responsive to local contexts and the necessity of changing harmful social norms (CI to UNAIDS 8/8/08).

Communication clearly has a key role to play and there are growing calls for broader social campaigns or social movements to address such issues.

These calls are far from new. Discussing successful HIV prevention efforts in 2004, David Wilson notes "above all… they were based on the premise that communities, however disparate, have within themselves the resources and capital to reverse this epidemic" (848).

In 2005 the CSIS Task Force (Kates & Nieburg) referred to the evaluation of HIV/AIDS prevention as an area "complex, often controversial, and one that will become increasingly important over the long haul" (2). They stress the need to focus on prevention and indicators of success, bemoaning the widespread neglect of HIV prevention targets compared to antiretroviral treatment goals. Like many they highlight that without greater prevention "there is significant risk of failing in efforts to stem the tide of the pandemic" (3). The key challenges to measuring HIV prevention they outline are all relevant to Social Change and SCC:

- Lack of a uniform definition of HIV prevention
- The challenge of measuring what did not happen
- Inevitable diversity and multiple models of success: "the few select cases of developing country success and even reversal… are due to multiple interventions and are themselves subject to much discussion and debate over the contributions of different factors. Prevention strategies, and their measures of success, will need to vary in different contexts, taking into account such factors as prevalence rates, "maturity" of the epidemic, the populations affected, and trends in new infections" (4).
- The time required to show impact; it takes years at the population level and interim, proxy measures are needed to capture progress towards (4).
- Urgent need to scale-up, with evidence of what works in prevention mainly demonstrated at individual, small group or community levels and not population (national) level, "because most HIV prevention interventions have not been implemented widely enough, or for long enough, to be able to show this level of impact" (4).
- The "complexity of behavior change and its measurement… since it is influenced by numerous individual and social factors" (4).

These pose significant challenges but do not justify complacency. It remains sadly true that a relatively small percent of AIDS funding goes to prevention, despite evidence of its impact and repeated calls for greater investment, at all levels.

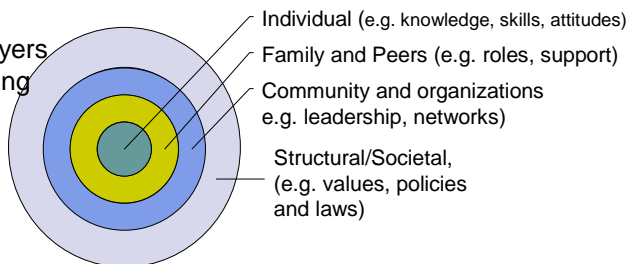**II. Characteristics of Social Change and the implications for its evaluation**
"Social Change" encompasses diverse multi-faceted, complex processes that are not easily-defined, but which share some key characteristics. The visuals below, from different contexts and presented in Kampala (SCC session slides), illustrate both the richness and complexity of comprehensive social change processes.

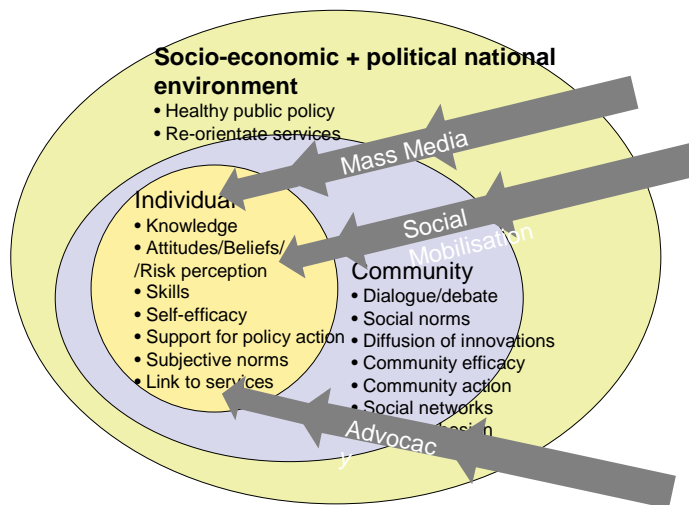## A *strategic, multi-level, multi-component* response is needed

- Social barriers (e.g. harmful gender norms) involve many reinforcing "parts," so change requires changing several things together (i.e. cumulative, sufficient, and moving in the same direction)

"Ecological" models illustrate "levels" or layers of influence surrounding individuals

Individual (e.g. knowledge, skills, attitudes)

Family and Peers (e.g. roles, support)

Community and organizations e.g. leadership, networks)

Structural/Societal, (e.g. values, policies and laws)

5

The Soul City Model of Social Change

Socio-economic + political national environment
- Healthy public policy
- Re-orientate services

Mass Media

Social Mobilisation

Individual
- Knowledge
- Attitudes/Beliefs/ /Risk perception
- Skills
- Self-efficacy
- Support for policy action
- Subjective norms
- Link to services

Community
- Dialogue/debate
- Social norms
- Diffusion of innovations
- Community efficacy
- Community action
- Social networks

Advocacy

14

Such models highlight the difficulty of attributing change to any one intervention, an area attracting increased interest, with good reason. As Iverson (IDRC, 2003) notes, "as sectors become more 'complex' in nature, attribution becomes more difficult. Similarly, as the level of intervention moves from the 'simple' project level to the 'comprehensive' program level, attributing results typically becomes

less feasible" (3). Likewise, discussing complex adaptive systems (CASs) which essentially is what social change encompasses, Eoyang & Berkas note "as the time horizon moves out, uncertainty increases because there will be more opportunities for various conditions and more time for those variations to magnify uncertainty".

The implications for evaluation include distinct strategies to assess a range of short-, medium- and long-term goals. In a SCC context short-term gains might include, for example, enhanced communication capacity or deeper dialogue, as well as the contribution this increased capacity or dialogue is making to broad, longer-term social change and more defined improvements relating to HIV/AIDS prevention and care. In such a context, *progress towards long term social change* and *the contribution being made* is increasingly considered a more realistic measure of effectiveness. This reflects a wider shift in the evaluation field, away from *measuring* and "*provin"*, towards *understanding* and *improving*.

The demonstration of SCC impact will never be quick or simple, as the Assessing Social Change (ASC) group at IDS highlights: "First, progress towards social justice and transforming relations of power does not follow a linear or predictable trajectory, with uncertainty beforehand about the impact and the most effective route. Second, the system-wide change that is being strived for requires efforts by and depends on multiple groups on diverse fronts; hence the utility of a focus on proving achievement in order to attribute impact to specific players is questionable. Third, drawing the lines in a process of social change with fuzzy and moving boundaries means valuing incremental shifts. Fourth, recognising a valid result requires valuing efforts. And finally, it is essential to acknowledge the timeframe of change (and divergence between realities and project lifespans) and, therefore, clarify expectations of change" Guijt, 2007: 5).

Learning from Complexity Theory and Systems Thinking
Complexity thinking encompasses an inclusive, overarching approach to understanding societies as complex systems and subsystems, all of which interact with each other. It involves combining and organizing observations and knowledge to bring together different disciplines and their contributions into a wider framework for understanding societal processes (Fowler, 2008).

One defining characteristic is overriding concern with social *processes* rather than social *structures.* Complexity theory "is precisely about understanding the tricky dynamics of achieving intended change… [and] does not assume that what is planned will necessarily occur" (Fowler). Rather, it focuses on what emerges from interactions. This is very significant in a SCC context: It highlights the links between context-specific social processes, standards, norms and values and, therefore, the danger of assuming replicability or scale-up.

Williams & Iman (2007) highlight the characteristics of complexity theory and systems thinking that share fundamentals with PM&E. These include:

1. A shift in focus to interrelationships and processes rather than snapshots, which seriously challenges dominant linear explanations of systemic phenomena.
2. An understanding of development as complex, emergent and transformative.
3. A shift to the bigger picture and interconnections, with much focus on boundaries and the values they reflect: "Boundaries are fundamentally about values – they are judgments about worth. Defining boundaries is an essential part of systems work/inquiry/thinking" (Williams, 2007). E.g. systems thinking raises questions about how you establish what is included or left out of the scope of work (i.e. which *interrelationships* were relevant), how you establish what standpoint to take (i.e. whose *perspectives* were relevant), and who or what (ideas) benefits or otherwise from those decisions (ibid).

Of particular relevance to social change and social change communication, Eoyang & Berkas' work usefully highlights how key features of CASs impact their evaluation. The essence of their argument is that "to be effective… an evaluation program must match the dynamics of the system to which it is applied" (1), as myself and many cited here have long advocated. Thus in contexts which include social change, "prediction and controlled performance toward a goal, like those assumed by traditional evaluation methods… cannot be expected from a complex adaptive system (CAS)" (2). Characteristic behaviors of CASs that need to be reflected in their evaluation are outlined below. They vividly illustrate the limitations and weakness of traditional evaluation paradigms.

Five main characteristics of CAS:
(i) Dynamic – Openness to external influences means a CAS is constantly changing, not in a smooth, predictable pattern but in… "jumps and discontinuities… (4).

Many others have made similar points about the need for flexibility, responsiveness and openness to what emerges: "The cyclical nature of planning, M&E is important because development programmes are part of an open system. The context in which a programme operates is continuously changing, so staff need to be engaged in ongoing reflection and learning so that the programme remains relevant and appropriate. (Earl & Carden, DinP 2002: 521).

(ii) Massively entangled – complex interrelationships mean incorporating multiple strategies, cycle times, time horizons, dimensions and informants: "Because a CAS has a structure that is non-linear, open and high-dimensional, an evaluation design cannot pre-determine all factors that will be of interest… it is critical that a variety of data be collected to reflect the variability of the system" (10). This can include linear evaluation strategies in short time frames and closed parts of the system, when appropriate. To give a concrete example, "adolescents' behavior may depend on age, gender, physical type, family stability, relationships with

peers and so on. In addition, over time, the relevance, power, or interrelationships among differences may change. Such a complex interaction of variables makes it unrealistic to expect to represent the system with a finite number of independent and dependent variables". Questions are raised about traditional evaluations that attempt to "identify a small number of key variables that affect change and to establish the relationships among those variables" (5) and that typically assume pre-determined and "unidirectional causality" that ignores the cross-causal nature of the system (6).

The "entanglement of variables in a CAS involves both their number and non-linearity… the system performs as a set of non-concentric, interlocking spheres of influence… Such messy and multi-level causal relationships determine the behavior in a CAS and make it difficult for the evaluator to establish clear units of analysis or lines of causality" (5).

(iii) *Transformative* – The CAS transforms and is transformed over time. Effective evaluations will respond by: making evaluation part of the intervention: "As a transforming feedback loop, assessment activities should enrich and enhance the intervention activities"; Involve as many members of the system as possible in the evaluation design. "Discussion of the design can be a time of tremendous learning and adaptation on the part of all stakeholders. With effective and continuing feedback loops, the whole system can co-evolve and adapt to the needs and aspirations of participants and the community" (10). There is a wealth of PM&E literature and experience that testifies to the benefits (Byrne, *et al*, Parks et al, Estrella and Gaventa).

(iv) Emergent – System-wide behaviours of a CAS emerge over time so it is important to track emergent patterns and pattern changes over time.

(v) Scale independent – It incorporates many levels and evaluation must incorporate both micro- and macro- patterns and structures.

There are significant implications. As Eoyang & Berkas state: "these principles, consistent with CAS behavior, shift the focus, tools and techniques of evaluation from the structures, low-dimension, predictable patterns of much of traditional research to more organic and flexible strategies. They also provide more structure and pre-designed rationality than many of the individualistic and constructivist methods of qualitative evaluation. By including a wide range of approaches, CAS methods of evaluation integrate the best of many disciplines and methods that were previously irreconcilable" (11).

Complexity theory clearly encompasses an understanding of change that privileges *networks, relationships* and *process.* As Eyben discusses in depth (see below) this is in stark contrast to the last three centuries in which "mainstream Western thought has understood historical change as linear progression (Eyben: 25).

As Iverson highlights: "within certain branches of evaluation research, methodological considerations have proven to be more acute—notably, those that deal with comprehensive interventions that are embedded in complex social systems [as in SC and SCC]. Perhaps nowhere is this more evident than in the field of international development research wherein the socio-economic, environmental, political and cultural dynamics of 'aid' efforts provide highly unique challenges for evaluators; where change is seldom attributable to any single factor, and can be extremely unpredictable. Moreover, faced with the external pressure to demonstrate results, development organizations and stakeholders are increasingly 'burdened' by having to prove the value of their initiatives to legitimize their own work. It is from this burden—from these challenges—that creative alternatives to the traditional modes of conducting evaluation research have evolved" (Iverson: 7). This is evident in much participatory research and evaluation, including innovative evaluation schools of thought as discussed below. As Bonbright concludes, "we are dealing with a high level of complexity and… in order to come up with feasible ways of measuring, we are going to have to reject a host of inappropriate concepts and tools" (1).

The relevance to Social Change Communication
In participatory development and communication "dialogic communication" is the norm: "Dialogue is needed in order to build trust, ensure mutual understanding, explore different perspectives and identify the best course of action to successfully address a situation that needs to be changed" (Mefalopulos: 2).

However, "how to measure the weight of such inputs is not an easy task, since before it reaches a stage in which it becomes visible and measurable, their impact is combined and diluted with a number of other variables" (ibid: 4). Thus assessing the value of participatory communication is not easy, "especially if the parameters for measuring impact are taken from the mainstream framework of quantitative methodologies" (ibid: 5). This is in contrast to communication interventions associated with the monologic mode which "can be assessed quite accurately, since they usually address changes in awareness, knowledge, attitude and, ultimately, in behaviors" stemming from an initial baseline (2). The latter is more familiar to many and sits more comfortably with typical organizational cultures and M&E systems.

Mefalopulos outlines three basic ways to assess participatory communication: anecdotal evidence, impressionistic evidence and "the costs of non-communication evidence". Anecdotal evidence is demonstrated by strong correlation or a causal link, not common to social change. "Impressionistic evidence" is most favoured by participatory perspectives as it focuses on success as perceived by the ultimate "beneficiaries". The latter, "costs of non-communication evidence", includes evidence of problems or failure resulting from a lack of dialogue with stakeholders. Questions are raised about what evidence is needed by whom and why.

III. Shifts in international development and the evaluation field that presenting new perspectives

Wider shifts in development

A more politicized understanding of development, as social change rather than as projects delivered through the mainstream development system, is challenging "received wisdoms of M&E… based on a different understanding of development itself as complex, emergent and transformative" (Guijt, 2007: 51). Various wider trends are fuelling greater interest in innovative ways of monitoring and evaluating social change. These include: growing experience with participatory approaches in large development organizations, wider understandings of accountability with increased focus on "more interactive accountabilities", the repoliticisation of development, and "looking beyond conventional tools and techniques to the experiences of social movements over the years" (e.g. Guijt, ASC: 5). Add to this growing interest in complexity theory and systems thinking, as well as increased calls to acknowledge the theory of change that a particular development initiative subscribes to and realize its implications, e.g., what constitutes "success" and how to assess it.

Broader understandings of development and social change reflect increased focus on relationships and networks of development actors, and appreciation of the practical relevance of epistemologies (what constitutes "knowledge" and whose knowledge counts). Such understandings foreground key development actors and processes that meaningfully engage key actors, *i.e.*, participatory ones. These methodologies are far from new and have long been advocated.

Arguably, it is senior stakeholders who have been most reluctant to grasp their implications, as dominant research and evaluation biases and practice suggest. Chambers' extensive work in this area highlights the crux of the issue: "Participatory professionalism challenges power. Much professionalism has been linear, standardized, top-down and patriarchal. Participation challenges patriarchy and the power and security of many teachers… The new participatory professionalism embraces self-critical reflection and learning, unlearning and unceasing personal and professional change" (2004: 7).

The growing shift towards appreciating the importance of relationships, inter-relationships and networks to successful development efforts is reflected in Eyben's valuable discussion on *Power, Mutual Accountability and Responsibility in the Practice of International Aid: A Relational Approach* (IDS, 2008). She considers the additional possibilities available when we "conceptualise aid as a field of interdependent and dynamic relations that are played out in the absence of pre-established consensus or shared vision concerning desired changes" (Summary).

This has direct relevance to notions of accountability, a tenet of evaluation. Eyben argues for a shift from the dominant *substantialist* perspective, where mutual accountability is about strengthening mechanisms for regulating behaviour between autonomous parties, to *relationalism.* Relationalism starts from the premise that social actors both shape and are shaped by their social relations (20).

The relevance to evaluation is profound: "whereas mutual accountability requires identifying specific power holders, diffuse or relational power links to ideas of mutual responsibility and the effect we have upon each other and the wider system" (Summary). However, this "more radical alternative framing of international aid" raises uncomfortable issues. Such a framework advocates a political rather than technical understanding of accountability, with a focus on struggle for voice and justice.

There have been significant shifts to develop and strengthen evaluation methodologies and methods cognisant of the above. These validate and legitimise new experts (in particular lay- as opposed to just professional knowledge), and new types of knowledge (stemming from lived experience, as well as that which is learned). There are major implications for evaluation methodologies, methods and evaluators themselves. As Bonbright concludes more broadly, "our profession [development] is getting serious and smarter about what we measure, but we are beginning to understand that we need to do this by *being inclusive in who does the measuring* (italics added)" (2007: 4).

As Picciotto argues: "A new paradigm has transformed the evaluation environment and raised the bar for the fledgling development evaluation profession. In the new century, development evaluation will have to be transformed and liberated from its current strictures to reach beyond aid, beyond projects, and beyond top-down approaches. It should be more comprehensive, more participatory, and better adapted to the felt needs of society" (Picciotto: 520).

What the evaluation field offers
These fundamental principles play out in multiple shapes and types of evaluation that essentially prioritise critical reflection, learning and improvement, through participatory processes. Strongly based on dialogue, these evaluation frameworks so have much in common with SC and SCC approaches. Overall we are speaking of what Schwandt terms a pedagogical approach to evaluation, i.e. "a process of teaching and learning about the deliberation of value; one that is encouraging and facilitative of critical reflection and self-transformation in conversation with others".

Such a perspective evaluation itself entails a teaching and learning process, one "more about learning than judging; more about participants becoming critically aware of their own positions on issues and developing an understanding and

appreciation of new and different perspectives. …This learning process is made possible by dialogues of several kinds" (Schwandt, 2001:232).

The notion of *critical* dialogue is highly significant: "Stakeholders must be active participants for responsive evaluation to succeed to cover and include various interests and values. Dialogue is central to its success, because stakeholders learn about the experiences and frustrations of others only through conversations (Abma: 34). As Schwandt says, "a private matter for each individual… we come to reasonable and just answers to questions of appropriate means and ends through dialogue and conversation with others. Consequently, this kind of evaluation is committed to the goals of participation, collaboration, and cooperation in the exploration of the evaluative imperative at the center of practice" (Schwandt, 2005:103). Patton speaks of "process use", or "using evaluation logic and processes to help people in programs and organizations to learn to think evaluatively… Learning how to think evaluatively is learning how to learn… Specific findings typically have a small window of relevance. In contrast, learning how to think and act evaluatively can have ongoing impact… (Patton, 1998:227).

Such approaches to evaluation are fundamentally educational, "less concerned with the provision and mastery of technical and narrowly instrumental knowledge of effectiveness and goal attainment, and more concerned with the lived practice of making evaluative judgments in specific situations" (Schwandt, 2001:233).

Like a relational approach discussed above, "Responsive evaluation" recognizes that the meaning of the world is shaped in interactions ". The goal of responsive evaluation is to enhance the understanding of a program from the life world perspective of insiders (*Verstehen),* and in that sense, the approach differs from evaluation approaches aimed at prediction and control" *(Erklaren*) (Abma:32).

Fundamentally such perspectives entail taking evaluation out of the realm of "experts" and capitalising on local knowledge, strengths, experience and abilities. As the participatory development guru Robert Chambers reflects, the notion that "they can do it" is the most powerful learning from his extensive experience: "local people… able or not able to read and write, were capable of complex mapping, diagramming and analysis to an extent that few if any of us 'professionals' had ever dreamt" (3). Such frameworks highlight the importance of relationships and evaluation grounded in practice aims "to illuminate and open to critical reflection the kind of knowledge that resides not in scientific statements of program outcomes and effects but in practice. Thus, the kinds of knowledge it is concerned with are located in lived action… and in relations" (Schwandt, 2005:102).

United under the umbrella of participatory approaches, PM&E, a significant and growing body of comparable methodologies that foreground communication and

dialogue highlight how evaluation can itself encompass a social change process and can be a stimulant and catalyst for social change.

At the very least, the evaluation of SC and SCC should be integral to the social change process itself.

As indicators remain the dominant mode of evaluation, it is important to spell out their limitations. These include:

- They simplify complex realities and often miss important information about changes including *why* they occurred.
- They do not effectively capture qualitative changes in people's lives and experience, which results in limited information that often misses great richness and depth that lies at the heart of SCC.
- Indicators often miss crucial information about broader institutional issues, partnerships and less tangible dimensions of programmes and their impact. This is particularly relevant to SCC initiatives which focus on the underlying societal factors that drive HIV and AIDS, including gender inequalities, stigma and rights violations.
- Basing evaluation on narrow (pre-determined) goals misses achievements that lie outside of these. They do not explain *un*anticipated changes, which might be most significant, which inhibits learning from the unexpected.
- Indicators are typically not used flexibly, creating a misfit between effectively capturing impact and gathering information needed, versus that obtained.
- There is a widespread, questionable focus on quantitative indicators which are more familiar to most and easier to collect and analyse.
- The learning potential of evaluation is often missed. Key questions are who defines indicators, whose information needs they meet and whose they neglect.
- Indicators have proven not to foster local ownership of process or findings. In practice they are often considered an unfortunate, unnecessary imposition of dubious value.
- Indicator-based monitoring is time consuming, costly and often of questionable value to projects and key stakeholders themselves.

This is not to imply that indicators do not have a place in evaluation, or that they cannot be meaningfully developed and used. The above weaknesses can be somewhat mitigated by *a participatory approach to indicator selection and use*.

However, I advocate a wider approach to evaluation. The development and evaluation paradigms discussed above have fuelled, and reflect, increased interest in the power of stories to capture diverse experiences and perceptions. As Abma says, stories "reveal the meaning and ambiguity of everyday situations and experiences, and as such, stories illuminate what really matters to stakeholders" (Abma:33). Perhaps the best-known methodology explicitly based on systematic story collection and analysis is the most significant change (MSC)

approach. This is attracting increased interest internationally and readers are encouraged to see Davies & Dart (2005), and Alliance India (2007) for its value in a HIV/AIDS context.

Guijt outlines the strengths of MSC, which "makes possible diversity of perspectives from diverse stakeholders, which might have prevented agreement around a single set of indicators. It acknowledges different sets of values between stakeholders and therefore does not force the use of tightly defined and static indicators. Contextually-explained stories that represent changes significant to the overarching goals of the intervention are valued above synthesized statistics. Broad categories of changes are identified by project staff but these are interpreted by the distinct groups. Recorded change events reflect a changing world and a changing set of perceptions about what is important" (Guijt, 2004).

Another methodology attracting increased interest is outcome mapping (OM), which is "premised on the belief that those engaged in the programme can benefit from working as a group to systematically collect, analyse and interpret data" (Earl & Carden, 2002: 522).

Like the frameworks cited above, OM explicitly prioritises learning by encouraging "evaluative thinking, participatory decision making, open sharing of successes and failures, and a willingness to engage in regular processes of thoughtful reflection and learning" (ibid: 519). It stems from critique of linear explanations of systemic phenomena. Rather, in line with systems thinking, OM recognises that a programme does not operate in isolation from wider external factors and therefore cannot plan or evaluate as though it did.

All involved need to reflect on what makes for sound evaluation in the social change and SCC fields. Heeding the oft repeated call for greater consistency between stated values and own practice, Eoyang and Berkas stress that learning should be the primary outcome as "effective adaptation is the best indicator of success in a CAS" (15). Programme contexts are ever-changing and "staff need to be engaged in ongoing reflection and learning so that the programme remains relevant and appropriate. It is impossible to plan for all eventualities; therefore, a successful programme is one that assesses and adapts to changing situations in a intelligent way based on thoughtful reflection" (Earl & Carden, 2002: 521). This is particularly true of social change and we are warned against being limited by preset performance objectives that can stifle vision and values and undermine openness to change and emergence (Patton, 2007: 114).

IV. How might we move forward in constructive, practical ways?
My intention has not been to overwhelm readers but to consider what we can learn from diverse sectors and disciplines, to better assess social change and social change communication as well as to support those facing this challenge.

That is, to ensure that people feel that evaluation of their efforts appropriately captures achievements and stimulates necessary learning and improvements.

Any participatory approach, by definition, is emergent and can always be strengthened. What is important is to make changes in the right direction and to be more realistic about what social change and its evaluation can achieve.

Here are some recommendations:
1. Generate practical ideas and share inspiring examples, we need to invest in:
    - Concrete efforts to systematize and review respective benefits and limitations of different grounded case studies
    - Training efforts for social change organisations on how to assess social change in ways that recognize core non-negotiable principles and purposes
    - Peer support opportunities to request and receive support with assessment and learning processes
    - Seeding experimentation and detailed documentation of processes (Guijt, 2007: 55).

2. Keep evaluation at arm's length from managers of programs and funding agencies, to ensure that eval contributes to accountability, learning and adaptability (Picciotto: 520). Clearly, "the nature of evaluation questions should drive the choice of methods rather than the other way round" (ibid).
To achieve credibility and effectiveness, evaluation processes should combine independent evaluation, self-evaluation, and full involvement of citizens—the intended beneficiaries of development programs" (Picciotto: 519). There are important capacity development and support implications.

3.   Combine methods as appropriate and question the bases on which particular methods are selected. Bonbright comments that current development thinking favours a "much friendlier pluralistic model in which qualitative, quantitative, perceptual and empirical data can be assembled into a comprehensible whole that still honours the complexity of social change (2007: 1). Iverson reflects on the significance: "By employing multiple methods evaluators do more than strengthen reliability, they broaden the depth and range of understanding of the changes associated with interventions. Stimulating this shift toward mix-method evaluation are philosophical changes in the purpose and function of evaluation. Picciotto (2007) and Prowse (2007) have highlighted the limitations of RCTs, still largely favoured by senior stakeholders, which are particularly evident in SC contexts. They are not appropriate for issues of efficiency, equity and quality impacts, which "are better captured by qualitative, participatory impact evaluations (Picciotto: 517).

4. Ensure flexibility in evaluation design and practice: "External assessment processes are often too rigid to understand the dynamics and processes that lead to mature and sustained social change" (Guijt, 2007: 9). This is part and

parcel of designing evaluation processes closest to where they will be used: "General rules and short lists of simple rules can be developed to apply to the entire system, but specific evaluation plans must be designed to meet the unique needs of the local context. Not only will the measures be more relevant and meaning, but the process of designing the evaluation plan will contribute to the system transformation process" (Eoyang & Berkas: 10).

5. Look actively for community strengths, assets and achievements, rather than being focused on problems and weaknesses: "Use evaluation as a reinforcing rather than dampening feedback mechanism, especially early on evaluation should reaffirm and amplify energy and commitment" (Eoyang & berkas: 10). This endeavour reflects the fundamental principles of participatory development and SC, and is fuelling growing interest in methodologies like Appreciative Inquiry.

6. Seek to redress imbalances. There is much talk of empowerment, participation, local ownership and equity in today's development discourse. Their implications for evaluation can be considered at different levels. From a global perspective, Picciotto discusses how striving for equity requires "principled partnerships that level the evaluation playing field, treat developing countries as full owners of evaluation methods and processes, and ensure fair and rigorous assessment of regional and global policies and programs" (Picciotto: 520).

At national and more local levels the implications are similar. A key strength of participatory approaches to evaluation is that they actively foster equity and voice, as well as collective reflection and learning for improvement. Evaluation capacity too often remains weak, whatever the methodology, and adequate resources and support are required, over time, to redress imbalances at all levels.

7. Donors should reconsider their evaluation policies and practice, as a recent expert group on assessing social change highlights. Practically, donors "need to rethink the principles on which they base their models of evaluation and learning", to ensure greater consistency between espoused values and the systems they use (Guijt, 2007: 52).

8. Intermediaries have key roles to play to bring about the changes advocated above. These include helping to ensure quality evaluations by developing capacity across the board, seeding experimentation, proposing alternatives to complement and strengthen current practice and "dialogue with donors to rethink the basis of assessment and learning processes" (ibid: 8). As Guijt highlights, intermediary roles focus on innovating, challenging, bridging and offering effective alternative ways of assessing and learning. This includes advocating for more open evaluation processes that help people engage without fear. It also includes fostering critical reflection on current practice.

The urgent need to strengthen the evaluation of processes of social change and social change communication is widely recognised. The implications are far-reaching and all development actors have a role to play. It means taking evaluation out of the realm of traditional "experts", with significant implications for capacity development and for support and resources required over time. It also calls for challenging dominant notions of validity, rigor and trustworthiness.

There is a long tradition of those opposed to change using rigour to critique participatory methods. It begs questions about which partner defines what "rigour" is, whether information can be "good enough to move forward" or must be "perfect" (according to whom?), and whose norms count (Guijt, 2000: 210). It brings us back to questions fundamental to any evaluation: Who is it for and why? Who can expect to gain what from the process?

We must remain open to learning from relevant experience and developments in other sectors and disciplines. In our increasingly interconnected world, links between (emergent) bodies of knowledge, and their mutual impact, serve to remind us that "we should not allow ourselves to become trapped by our disciplinary assumptions, specialist knowledge, or habitual approaches. We should have the courage to be creative in the way we think about explaining and promoting change" (Krznaric: 46).

Powerful forces continue to preserve the status quo and to undermine alternatives to assessing social change, implicitly and explicitly. This is regardless of issues of appropriateness and quality, never mind those of equity, voice and value to particular groups and contexts. Do not underestimate the influence of the dominant paradigm and evaluation "machinery" that is witnessed in legitimised experts, capacity development and training opportunities and the resources available or not for particular evaluation methodologies and methods. As Robert Chambers has repeatedly emphasised, questions must be asked about *un*learning and *re*learning required, and about who needs to learn. The importance of openness to change and to a degree of risk-taking cannot be overstated.

I will conclude with reflections made recently by Michael Patton, following his decades of experience and leadership in the field: "International diversity is challenging our thinking about what constitutes good evaluation work and what it means for evaluation to be used in different cultural and political contexts… That narrow form of defining what "true" evaluation is… is challenged by the different cultural and political ways people think about knowledge, what constitutes knowledge, what constitutes evidence, how evidence impacts a political context, and the dramatically different role of nonprofits and governments in different places" (Patton, 2007:112).

Often the most difficult, challenging experiences result in the greatest learning. For those of us working in this field, we are constantly challenged, stimulated and inspired and the process is rich in learning.

We welcome the constructive engagement of more actors in this effort.

Works cited

Abma, T (2006). *The Practice and Politics of Responsive Evaluation.* <u>American Journal of Evaluation</u> Vol. 27(1): 31-43

AIDSCAP (1997). *Making Prevention Work. Global Lessons Learned from the AIDS Control and Prevention (AIDSCAP) Project 1991-1997.* FHI/USAID

Alliance India (The India HIV/AIDS Alliance) (2007). <u>Stories of Significance: Redefining Change. As assortment of community voices and articulations.</u> A report based on an evaluation of a programme on "Community Driven Approaches to Address the Feminisation of HIV/AIDS in India" by means of the 'Most Significant Change' Technique. http://www.aidsallianceindia.net/Publications/Stories_of_Significance_Redefining_Change.pdf

Bonbright, D (Dec 2007). Measuring Impact? Who counts? Making Social Investment Decisions, What do we need to know? <u>Alliance Magazine</u>. London: Keystone Reporting. http://www.alliancemagazine.org/free/html/dec07e.html

Byrne, A (2007). *Working Towards Evidence-Based Process: Evaluation That Matters.* MAZI 13 (November 2007). New Jersey: CFSC Consortium

Byrne, A, (Ed) with Gray-Felder, D, Hunt, J & Parks, W (2005). *Measuring Change: A Guide to Participatory Monitoring and Evaluation of Communication for Social Change.* New Jersey: CFSC Consortium http://www.communicationforsocialchange.org/pdf/measuring_change.pdf

Byrne, A, Gray-Felder, D & Hunt, J (Eds) (2005a). *Communities Measure Change. A Reference Guide to Monitoring Communication for Social Change.* New Jersey: CFSC Consortium http://www.communicationforsocialchange.org/pdf/communities_measure_change.pdf

Chambers, R (2004). *Reflections and directions: a personal note.* <u>Participatory Learning and Action</u> 50, Oct 2004. London: IIED

Davies, R & Dart, J (2005). The 'Most Significant Change (MSC) Technique. A guide to its use. http://www.mande.co.uk/docs/MSCGuide.pdf

Earl, S & Carden, F (2002). *Learning from Complexity: the International Development Research Centre's experience with Outcome Mapping.* <u>Development in Practice</u> Vol. 12(3&4): 518-524

Eoyang, G & Berkas, T (April 30, 1998). *Evaluation in a Complex Adaptive System. http://www.chaos-limited.com/EvalinCAS.pdf* (accessed 7 Nov. 2008)

Eyben, R)May 2008). 'Power, Mutual Accountability and Responsibility in the Practice of International Aid: A Relational Approach'. IDS Working Paper 305. Sussex: IDS.

Fowler, A (2008). Connecting the dots: Complexity thinking and social development. The Broker, 7, April 2008.

Guijt, I (2000). Methodological issues in participatory monitoring and evaluation. In Estrella, M et al (Eds), Learning from Change. Issues and experiences in participatory monitoring and evaluation. London: Intermediate Technology Pubs: 201-216.

Guijt, I (2007). Assessing and Learning from Social Change: A Discussion Paper. Brighton: IDS.
http://www.ids.ac.uk/UserFiles/File/participation_team/publications/ASC_low-res_final_version.pdf

Guijt, I (2008). *Critical Readings on Assessing and Learning for Social Change: a literature review.* IDS Development Bibliography 21. Brighton: IDS.
http://www.ids.ac.uk/UserFiles/File/participation_team/Db21.pdf

Inagaki, N (2007). *Communicating the Impact of Communication for Development. Recent trends in empirical research.* World Bank Working Paper series, Development Communication Division. Produced for the World Congress on Communication for Development in Rome, October 2006.
http://www.comminit.com/en/node/266501/36

Iverson, A (2003). *Attribution and Aid Evaluation in International Development: A Literature Review.* For Evaluation Unit, IDRC, Canada.

Kates, J. and Nieburg, P. (2005). Prevention indicators for the President's Emergency Plan for AIDS Relief. A report of the CSIS Task Force on HIV/AIDS Working Committee on Prevention. Washington, D.C.: Center for Strategic and International Studies.

Krznaric, R (2007). How Change Happens. Interdisciplinary Perspectives for Human Development. Oxfam Research Report. UK: Oxfam

Lacayo, V (2008). *When It Comes to Social Change, the Machine Metaphor Has Limits.* MAZI 16, August 2008. New Jersey: CFSC Consortium

Lamptey, P & Cates, W (1996). AIDScaptions, Vol III, No.3. The Case for Prevention. FHI.
http://www.fhi.org/en/HIVAIDS/pub/Archive/articles/AIDScaptions/volume3no3/OzPrevention.htm

Long, N. (2002). *An actor-oriented Approach to Development Interventions.* Background paper prepared fro APO Meeting, Tokyo 22-26 April 2002.

Mayne, J (June 1999). Assessing Attribution Through Contribution Analysis: Using Performance Measures Sensibly. Discussion Paper. Canada: Office of the Auditor General of Canada. http://www.oag-bvg.gc.ca/internet/docs/99dp1_e.pdf

Mefalopulos, P (2007). *Participatory Communication. When the beginning defines the ending.* Glocal Times 23/11/07. Sweden: Malmo University: http://www.glocaltimes.k3.mah.se/viewarticle.aspx?articleID=123&issueID=15

Parks, W, Gray-Felder, D, Hunt, J & Byrne, A (2005). *Who Measures Change? An Introduction to Participatory Monitoring and Evaluation of Communication for Social Change.* New Jersey: CFSC Consortium http://www.communicationforsocialchange.org/pdf/who_measures_change.pdf

Patton, MQ (1998). *Discovering Process Use.* Evaluation Vol. 4(2): 225-233.

Patton, MQ/ Oral History Project Team (2007). "The Oral History of Evaluation, Part 5. An Interview with Michael Quinn Patton". American Journal of Evaluation Vol. 28(1): 102-114

Picciotto, Robert (2007). *The New Environment for Development Evaluation.* American Journal of Evaluation Vol. 28, No. 4: 509-521.

Prowse, M (Dec 2007). *Aid effectiveness: the role of qualitative research in impact evaluation.* Background Note. London: ODI. http://www.odi.org.uk/events/development_evaluation/ODI-V3.pdf

Reilly, M (Feb 2007). An Agenda for Change in the USA: Insights from a conversation about Assessing Social Change in Washington, DC. Washington: Just Associates. Case study for the Assessing Social Change project at IDS. See Guijt (2008). http://www.ids.ac.uk/UserFiles/File/Participation_publications/JASS_casestudy.pdf

Schwandt, T (2001). *Understanding Dialogue as Practice.* Evaluation Vol. 7(2): 228-237.

Schwandt, T (2003). *Back to the Rough Ground! Beyond Theory to Practice in Evaluation.* Evaluation Vol. 9(3): 353-364.

Schwandt, T (2005). *The Centrality of Practice to Evaluation.* American Journal of Evaluation. Vol. 26(1): 95-105.

UNAIDS (May 2008 draft, soon to be published). 'Technical Update on Social Change Communication. A document outlining the rationale, principles and steps for social change communication. Geneva: UNAIDS. Report of the UNAIDS Technical Consultation on Social Change Communication.

Williams, B & Iman, I. (Eds.), (2007). Systems Concepts in Evaluation: An expert anthology. CA: EdgePress of Inverness, pubs.

Wilson, D ((2004). *Partner reduction and the prevention of HIV/AIDS. The most effective strategies come from within communities.* Editorial. BMJ Vol 328 (April 2004): 848-849.