"Show me your impact": Challenges and Prospects for Evaluating Transitional Justice

Colleen Duggan

International Development Research Centre
Evaluation Unit

cduggan@idrc.ca

Working Draft – Please do not cite without permission from author

As the use of mechanisms for transitional justice[1] has proliferated throughout the world

over the last decade, but especially over the last five years, those involved in the business of

international aid are coming to realize that the record for evaluating the effects that these

interventions are having on the lives of people has not matched enthusiasm for promoting,

designing and financing them.

Transitional justice has come to occupy a position of increasing importance in the

administration of international aid;[2] so too has the menu of mechanisms or interventions called

for by international organizations, donor agencies, scholars and civil society advocates. Although

there is no established model for transitional justice – indeed, in 2004 a UN report warned that we

must "eschew one size-fits-all formulas and the importation of foreign models" (UN, 1) –

transitional justice has moved beyond the realm of the juridical to embrace a suite of mechanisms

that include:

(1) Criminal prosecutions of individual perpetrators (either through national or international tribunals);
(2) truth seeking initiatives to address past abuse and clarify historical accounts of the past (truth commissions, investigative bodies);
(3) material and/or non-material reparations to victims of human rights violations.
(4) commemoration for victims;
(5) reforms to key institutions (including the judiciary, army and police and vetting, dismissals and barring from public office);[3] and most recently
(6) reforms to history education.[4]

---

[1] The notion of transitional justice was coined in the mid-nineties (although transitional justice and experiences pre-date the nineties) and refers to how societies moving from repressive rule or armed conflict come to terms with large-scale past legacies of human rights abuse in order to ensure accountability, seek "reconciliation", examine historic memory and create justice systems so as to prevent future human rights atrocities. (Charles Call,"Is Transitional Justice Really Just?," *Brown Journal of World Affairs, 11(1)*: 101-113; United Nations Security Council. The rule of law and transitional justice in conflict and post-conflict societies. Report by the Secretary General. 23 August 2204. S/2004/616). Countries in political transition are increasingly considering both judicial and non-judicial responses to human rights crimes with differing levels of international involvement (or none at all).

[2] Footnote levels of funding from Louis` Bickford`s recent report for Ford Foundation here – need to ask permission.

[3] See generally, Louis Bickford, "Transitional Justice" in *The Encycopedia of Genocide and Crimes Against Humanity* (New York: MacMillan Reference USA, 2004), vol. 3, 1045 and International Center for Transitional Justice website.

[4] Elizabeth A. Cole makes a convincing case for including history education as one of the institutions that should appear in frameworks for transitional justice. "Transitional Justice and the Reform of History

As a field, it would seem that transitional justice is now coming of age. This paper is motivated by three recent trends that are influencing academic and policy debates about the merits and pitfalls of transitional justice. The first relates to the social science literature on transitional justice, which is calling into question some of the core assumptions that tie together transitional justice mechanisms and their potential to identify causal mechanisms upon wider processes for social change, including (but not restricted to) peacebuilding, reconciliation or `social healing`, democratization and consolidation of the rule of law.[5] The fact of the matter is that those with a stake in the promotion of transitional justice have been weak in consolidating a compelling and rigorous body of theoretical and empiric evidence that assesses both the positive and negative effects of these mechanisms upon larger processes of peacebuilding and development. Scholars who are interested in problematizing transitional justice point out that what we have is largely anecdotal or inconclusive evidence.[6]

In parallel to this debate, within the international aid community, there is a re-kindling of discussions around the effectiveness of aid and impact evaluation (and what constitutes acceptable or credible evidence). This manifests itself in one of two ways: A continuing tendency by donors to cling to linear, ill-adapted methods and approaches to the evaluation of aid for international development, despite a tacit (or explicit) recognition of their inadequacies[7]; or seizing upon the use of experimental and quasi-experimental methods as a `gold standard` for impact evaluation, with little consideration for the need for pluralist approaches that account for

---

Education" in *International Journal of Transitional Justice*, Vol. 1, 2007. In the latter part of this paper, I will examine a case study that focuses on history education in Guatemala.

[5] See for example Mendelhoff on truth telling 2004; Brahm on Truth Commissions, 2006; Vinjamuri and Snyder on criminal trials, 2004).

[6] For a thorough review of some of the major studies to date on truth commissions, criminal trials and vetting see Thoms et al, `*The Effects of Transitional Justice Mechanisms. A Summary of Empiric Research Findings and Implications for Analysts and Practitioners*` (upcoming)

[7] See for example Oliver Bakewell and Anne Garbutt. *The Use and Abuse of the Logical Framework Approach.*. SIDA, Stockholm, *2005; United Nations General Assembly. Review of Results-Based Management at the United Nations*. 22 September 2008. A/63/268. The by-line of this report is as follows: "Results-based management at the United Nations has been an administrative chore of little value to accountability and decision-making."

the many, complex factors that actually lead to change in transitional societies.[8] As both a researcher and a donor, I have seen these dramas being played out in major international organisations working in human rights and transitional justice in both the global north and south.

The third significant development that underpins this paper is the (not surprising) growing interest among researchers working on transitional justice to make use of the principles and methods emerging from international development evaluation (most commonly known as `program evaluation) to improve applied research on transitional justice.[9] Social science researchers and advocates are increasingly looking to program evaluation for new learning and clues that might help lead them out of the impact quagmire.

Moving to the level of practice, civil society organisations and applied researchers – particularly those working with victims in the south – are demanding legitimate approaches to better understand if and how transitional justice processes are effecting positive or negative change in their societies. International bilateral and multilateral donors and philanthropic grant-makers are evidently also interested in knowing if and how their investments are having `impact` - often for different and sometimes less altruistic reasons. Clearly, the single biggest dilemma facing researchers and practitioners working in the area of transitional justice is the urgent need to build an evidence base to underpin policy and practice. The clamour for more evidence-based findings on the `impacts` of transitional justice and the desire to advance our understanding around how international aid for transitional justice can best be evaluated – offer new opportunities and dilemmas for anyone with a stake in the success or failure of transitional justice.

---

[8] Renewed interest in such methods was kindled by the Centre for Global Development`s report, *When will we ever learn? Improving lives through impact evaluation*, Report of the Evaluation Gap Working Group, May 2006. Since the release of the CGD report, debates for and against expermental methods, especially randonmized controls trials have continued to rage. See Martin Ravallion, *Should the Randomistas Rule?*, Economists Voice, Feb. 2009, www.bepress.com/ev
[9] See Phuong Pham and Patrick Vinck, "Empirical Research and the Development and Assessment of Transitional Justice Mechanisms," *International Journal of Transitional Justice*, July 2007; 1: 231 - 248. 2007.

In this paper, I would like to explore some of the challenges and opportunities for evaluating transitional justice. The discussion will begin with an examination of why transitional justice is so hard to `measure` and why traditional approaches to program evaluation have so far, been woefully ill-equipped for advancing our understanding of the effects or `impacts` of these mechanisms. Using a case study, in the second part of this paper I will then empirically ground these challenges and share learning from the International Development Research Centre emerging from our efforts to track and evaluate the effects of a museum exposition that is attempting to recast historic memory and challenge racism in post-conflict Guatemala.

1.   Why is transitional justice so difficult to evaluate?

Prior to entering launching into the heart of this discussion, some conceptual clarifications are in order. When we talk about `evaluating` transitional justice` we are usually talking about two things: Using social science research (usually applied, often empiric) to evaluate the effects that transitional justice mechanisms have on the people and environments in which they are active; and evaluating the efficiency, effectiveness and relevance of the actual programs that act as vehicles for funding transitional justice mechanisms. Although this paper focuses on the evaluation of international aid programs in support of transitional justice, it is important to note that program evaluation and applied research have many similarities and a few differences. (See graphic 1).
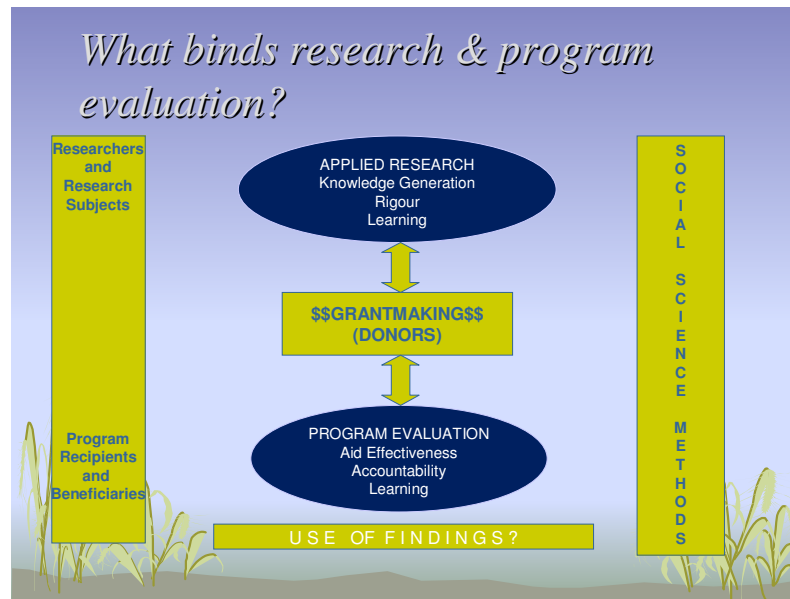
Graphic 1: Presentation to the Board of the International Journal of Transitional Justice.
Colleen Duggan, Capetown, May 2008

Both evaluation and applied research rely on social science methods and examine

multiple facets of a problem, often using multi-method approaches; both collect and analyse data

in order to come to conclusions and both utilize theory to inform work. However, the core

concept of program evaluation is its means of assessing the value or worth of something against a

set of criteria. Evaluation always assesses the performance of the program or organisation under

investigation. Another difference is audience: Program evaluation always has a client who wants

to know something. Good evaluation always uses universally accepted standards (utility,

propriety, feasibility and accuracy)[10]. So, in very broad terms, the major distinguishing

characteristic is judgement – evidence is used to judge merit or worth and this forms the basis of

decisions. In the world of international development aid, applied research and program evaluation

are tied together by research grantmakers or donors. Donors use program evaluation to find out

`what works` and make real-time programming decisions about the sorts of transitional justice

policies and programs that they will fund. Of course the ideal scenario would be to ensure that

---

[10] American Evaluation Association. *Program Evaluation Standards*.

program evaluation - which generally has tended to assess short-term measurable successes - is coupled with longer term empiric research which would constitute a more knowledge-based approach to the planning and administration of international aid. Indeed, in the field of transitional justice, there are those who would argue that much of the research on transitional justice mechanisms is still inconclusive and decision-making is based on an evidence base that is flimsy at best.[11]

Evaluation as a field or discipline is often maligned or frequently misunderstood, especially by those who have come to understand it as an unpleasant, bean-counting sort of exercise that is undertaken by technocrats for the principal or sole purpose of ensuring upward accountability between implementers and donors for funds spent. The truth of the matter is somewhat different: Evaluation is as varied in its methods and ideologies as is social science – it can be empowering or punitive; it can be qualitative, quantitative or mixed method. It can focus on gathering data on tangible results or outputs or process results. There is no perfect evaluation model or approach for transitional justice – only choices that need to be made. I will return to this point later.

In the field of transitional justice, the `field` has grown at neck break speed and the donor landscape has become dense indeed. In the last ten years or so, there has been increasing trend among bilateral and multilateral development assistance donors to fund transitional justice programs, including research for transitional justice. Arguably, an increase in funding opportunities through these donors is a good thing; the major difficulty, however, is that many of these donors are stuck in old paradigms of program evaluation. This is having a profound effect not only upon the administration of international aid but also upon how transitional justice research is perceived and `measured` and what counts as evidence.

Evaluating the effects or "impacts" of transitional justice faces a number of challenges of methodological, contextual and political dilemmas. I will speak to five of them.

---

[11] Thoms et al. *Op cit.*

Dilemma 1: Dealing with complexity and unpredictability

Transitional Justice processes most often take place in socially and politically fragile environments that are characterized by high and increasing complexity. Transitional justice mechanisms, especially when used in combination, aspire to facilitate or promote complex social change processes that require multifaceted interventions and multiple national and international actors over whom we do not have uniform levels of influence or control.

Typical change processes or implicit (and often untested) assumptions about the goals of transitional justice include social healing (through truth-telling) or facilitating the formation of new identities (through memorialisation or educational reform), to name a few. In new or recovering democracies these are long term processes that are cyclical in nature, whereas transitional justice projects and programs funded by international donors tend to be shorter in duration – maybe three to five years (the duration of your typical truth commission, for example). The sort of impacts hoped for (e.g. rebuilding civic trust; building a human rights culture, creating empathy between former adversaries, etc.) are the results of years and years of investment and can take generations.

Transitional justice settings are most often dynamic, fluid and prone to relapses into violence. They are also highly sensitive. Impact is often linked or tied into the perceived authenticity of human relationships. Often transitional justice advocates themselves are accountable to their own constituencies in country or arguably outside or universally, as in the case of international lawyers, to cite one example. The roles and reputations of both international and local transitional justice advocates are not always clear; important factors such as integrity, impartiality and credibility are difficult to track, let alone "measure".

Dilemma 2: The problems of causality and attribution of results

Both the funders and the funded, face a political and operational conundrum that is inherent in international development: Attribution of results or change to a single or a series of

interventions is something that all donors dream of but rarely are able to collect enough evidence to prove. Changes in the well-being of intended beneficiaries of transitional justice can occur before or after a program ends; they may not take the form anticipated; and they may be influenced by the actions of stakeholders who remain beyond the reach of the program[12] (peace process `spoilers` come immediately to mind).

Proving causality and attribution of results is even more difficult in complex environments. And yet, this is exactly what most donors ask their program recipients to do. The attribution obsession in international aid has lead to a tendency to rush towards what I would argue is an exaggerated focus on results and a naïve belief that the existence of rigorous evidence is a sufficient guarantee to influencing policy. Evidence-based policy making seems to have become the new silver bullet of international aid and transitional justice advocates and researchers appear to be falling into the trap of believing that anything is provable – if only we have the data to back it up. The unfortunate reality is that evidence plays a small role in policymaking for development assistance and it must jostle for position amidst other considerations such as political imperatives and human relationships.

Dilemma 3: Linear vs. systems approaches to planning, monitoring and evaluation TJ

As concerns the management of aid for development, including transitional justice, it has played itself out by dividing evaluation into two camps with very different arguments about how to conceptualize and operationalise issues of impact and change: Linear approaches and systems approaches.

The first group can roughly be categorized by those who favour monitoring and evaluation systems that are based on linear cause-effect thinking, or causal chains. Programs and projects are generally laid out in logical frameworks (or log frames – those 6 X 8 tables you see

---

[12]Terry Smutylo. "Outcome Mapping: A method for tracking behavioural changes in development programs". *Institutional Learning and Change* (ILAC) Brief 7, August 2007.

that list objectives, inputs, activities, short and long-term outcomes, impacts). The underlying assumptions imbedded in this results-based management approach are that we know and can measure impact and progress through objective variables and we can, to a reasonable degree, predict the impact of our programs during the design and implementation phases. [13]

The second group are characterized by their preference to work within a systems or complexity approach. This group approaches peacebuilding and development through a more elliptical lens; they are relationship focussed and have a desire to be flexible and responsive to each situation. This group generally does not believe that events in conflict or transitional environments can be predicted because they are part of a larger complex system made up of intermeshed forces over which we as transitional justice advocates or development agents have little (if any) influence. The assumptions that underpin this approach include the belief that every situation is unique, lessons are not transferable from one country or setting to another, planning has limitations and flexibility is always an asset. [14]

Logical Framework approaches have their supporters and their detractors; in fact, log frames, when developed and used in participatory and dynamic ways can be very powerful tools for strategic thinking. The problem is that logical frameworks have been abused far too often over the years. Time-strapped donors like them because they apparently simplify – in a very misleading way - what are usually very complex situations in which change is always multi-causal. The other difficulty with linear approaches is that in dynamic, fast-moving and politically fragile environments, they do not lend themselves well to adaptation - hence their nickname, `lock frames. Transitional justice as a field is emergent field.  We still have little if any evidence that TJ actually `works` and that lessons can be compared across contexts. This suggests that our approach to evaluation needs to make ample room for innovation and adaptation If we are to

---

[13]Reina C. Neufeldt. " `Frameworkers` and `Circlers` - Exploring Assumptions in Peace and Conflict Impact Assessment". Berghof Research Center for Constructive Conflict Management. August 2007. http://www.berghoff-handbook.net, Accessed October 7, 2007.
[14] Neufeldt, *op cit.*

ensure that the people actually being affected by transitional justice policies are not to be permanently entrenched as the new guineas pigs of international aid we need to ensure that there is ample opportunity for program learning, especially at the local level.

The other difficulties with linear approaches spring from practice. Too often log frames have been imposed upon program recipients with little or minimal training and support in their use. Fund recipients tend to fill them in as a donor requirement with little regard for their use or their updating. This is problematic for many reasons, but three in particular stand out: First, it encourages recipients to exaggerate intended outcomes in the hope of securing funds. This is problematic for transitional justice processes because often times the ultimate beneficiaries of program funds are victims; Second, it puts the donor-recipient relationship on unhealthy footing and exacerbates power differentials that are often already acute (a point I will come back to further on) and; third, poorly conceptualized and under-utilized log frames put recipients (and donors) in difficult circumstances when program outcomes fall far short of what was predicted at program inception – and this is always the case in complex social change environments.

Linear approaches to evaluation also often unwittingly stifle innovation and penalize adaptation. This is highly problematic in contexts in which learning and adaptation needs to be a top priority for victims, for victims support groups, for governments and all other stakeholders - including the donors themselves – who are invested in transitional justice outcomes. The use of such approaches in sensitive post-conflict or post authoritarian situations that are highly politicized and generally do not lend themselves to replication - has cultivated a false sense of being able to control how change happens. This is a slippery slope because it belies or underestimates the importance of agency. In addition, the method and its purported infallibility, is starting to become an end unto itself.

Dilemma 4: Dealing with trauma and power differentials

All the literature on transitional justice and indeed our own donor rhetoric states that relationship building is at the heart of transitional justice approaches in support of reconciliation and democratic strengthening. Improving trust in strained inter-group relationships should be seen as a priority. Add to this, the fact that we are usually working with highly traumatized populations; whatever form of monitoring and evaluation we use, it should not and cannot re-victimize populations or make an already fraught situation worse.

And yet in many cases, the role that evaluation plays in these contexts is far from supportive. Unfortunately, evaluation is still largely viewed by program recipients as a negative and often punitive experience that imposed from on-high. In a workshop of African transitional justice researchers and practitioners supported by the IDRC, most participants expressed great scepticism about the indicators, evaluation methodologies and set of values that tend to accompany external evaluations lead usually by northern trained "experts" who have been parachuted in to undertake an independent evaluation. Typical comments were:

"Evaluators tend not to ask critical questions about processes and relationships."
"Evaluators tend to focus on outputs not outcomes."
"Evaluation loses credibility where there is limited joint dialogue, iteration and feedback."[15]

On the other hand, the same participants were optimistic about the opportunities for improving evaluation of transitional justice and better understanding changes and impact. They said:

"There is an opportunity to improve/adapt evaluation techniques in rapidly changing societies."
"Ownership by the person or entity being evaluated improves the quality of the evaluation."
"Political will exists but implementing organizations require capacity and support to engage in the evaluative process."[16]

Evaluation of programs designed to serve disadvantaged groups such as victims may actually be threatening to stakeholders in those groups. Program beneficiaries, like research subjects – for often they are the same – have also suffered bad experiences with evaluation. For

---

[15] International Development Research Centre-Center for the Study of Violence and Reconciliation. "Evaluating Experiences in Transitional Justice and Reconciliation: Challenges and Opportunities for Advancing the Field," Workshop Report. Ottawa: CSVR-IDRC, April 2007 at 18.
[16] IDRC-CSVR, *op cit* at 18.

most of these people, evaluation as been a highly disempowering experience. Evaluation like applied research is an external intervention – in the case of program evaluation, the power differential is huge, considering that beneficiaries often know or presume the element of judgment and consequences for future funding and livelihoods.

Because such a large part (although not all) of transitional justice is about re-building broken relationships, minimally, the evaluation approaches used in these contexts need to focus on: Empowerment of victims, international and local stakeholder accountability and recuperating or reconstructing lost social capital. Many of the goals of transitional justice suggest that participatory evaluation approaches such as self-assessment and peer review - often akin to action research methods could play a constructive role in all of these processes. When coupled with external evaluation, these methods could underpin emerging evidence on results about the effectiveness of transitional justice with equally important considerations of legitimacy and ownership of the process as a whole. In addition, it would build learning capacities into the organisations of transitional justice stakeholders. At the end of the day, it is local actors who need to be convinced that transitional justice `works`. Uptake and ownership of the well-intentioned goals of transitional justice will depend upon whether local actors perceive the achievement of results that are of import to them and their constituencies – not the international donors who fund these processes and their evaluators. Unfortunately, program accountability too often means vertical exclusive `accountability` to funders in distant capitals rather than downward and horizontal accountability on the ground.

Dilemma 5: Unclear or inexistent theories of change

While a number of scholars are now complaining about the grandiose untested assumptions of cause and effect that tend to accompany transitional justice, I would put to you that the problem goes deeper than sloppy science. Because administrators of international aid tend to be governed by time-bound programming cycles, they often are unable to see beyond the

myopic, short term output-based results that lie within the sphere of the immediate project they are funding. This `project trap` inhibits any possibilities for both the funders and the funded to come to grips with the fact that transitional justice, like many interventions for peacebuilding and governance are highly dependant upon often implicit assumptions about how long term change happens. Too often international aid for transitional justice interventions has been devoid of any well articulated theory of change – a process that explicitly lays out a causal chain and maps out the assumptions that guide project design at inception to the long-term goals it seeks to achieve. While this may sound simplistic to any social science researcher, it is actually a perennial problem in funding for complex development interventions and is particularly important in cases in which `soft` or `fuzzy outcomes` such as `increased voice for victims` or `enhanced accountability` or `social healing.` are not easily identifiable or verifiable.

2. **Historic Memory and Racism in Guatemala: Evaluating the Effects of the "*Why are we the way we are*?" Museum Exposition**

How the story begins

The Guatemalan internal armed conflict ended in 1996, leaving about 2% of the national population dead or disappeared – 83% of who were indigenous.[17] In 1999, Guatemala`s Truth Commission concluded that the Guatemalan state had committed acts of ethnic genocide against the indigenous population, placing Guatemala among a small group of nations that face the challenge of healing a society, which allowed the "unthinkable" to happen. Although most Guatemalans recognize that racism is a problem, constructing a nation in which ethnic diversity is celebrated and everyone is a citizen in the fullest sense of the word is a task for many generations to come.[18]

---

[17] Insert reference, Comision de Esclareciminiento Historico, 1999.
[18] Tani Adams, *Proposal to Develop an Educational Strategy or Teaching about Racism and Ethnic Equality in Guatemala: Por Qué Estamos Como Estamos? 2007- 2011*, 2007.

It was in this context that the Centro de Investigaciones Regionales de Meso América (CIRMA,) a research centre and library founded in the midst of the armed conflict in 1980, decided to dedicate major attention to developing research and educational strategies around the issue of racism, as a contribution to the social reconstruction of Guatemala in the post-war period.[19] IDRC`s relationship with CIRMA actually pre-dated the release of the Truth Commission Report. Although the issue of history education did not figure prominently in the CEH report, interethnic relations and the transformation of Guatemala`s education system was clearly recognized as a priority in the Peace Accords.[20] Between 1998 and 2003, IDRC, through its Peace, Conflict and Development Program (PCD) had supported CIRMA to produce an exhaustive multi-volume ethnographic study of the history and current tendencies of ethnic relations in Guatemala.

In 2003 CIRMA began discussing with IDRC and other donors the potential for using this research to reconstruct a new Guatemalan narrative around race relations and the history of the conflict. CIRMA`s research formed the basis for an ambitious nation-wide effort to stimulate public reflection and dialogue, called the National Campaign for Interethnic Dialogue: *Our Diversity is Our Strength!* The campaign was developed on the basis of a national consultation, which detected strong interest in addressing the issue of racism in Guatemala, and just as strong a fear as to how to address the issue without exacerbating existing tensions.[21]

This challenge – how to create a mechanism for dialogue which would foster, and not hinder, intercultural reconciliation – was at the heart of the design of the Campaign launched in June 2004. The Campaign had three components, developed with help from international experts on social violence and reconciliation and based on learning emerging from museums of conscience and race consciousness training experiences:

---

[19] Adams, *op cit.*
[20] See Agreement on Identity and Rights of Indigenous Peoples (1995). Need full cite.
[21] Adams, *op cit.*

o A 5,000 square foot interactive exposition, *Por Qué Estamos Como Estamos?*
(Why are we the way we are?) The expo presented the history and current state of interethnic relations in Guatemala in an animated and colourful way through photographic images, videos and interactive games that invite people to reflect on their own lives in Guatemalan society.

o A guided process for group dialogue on the topic: the dialogues groups were meant to accompany the expo and aimed to engage community leaders who were seen as multipliers in a discussion on how to transform the prevailing system; and

o A series of university teaching activities oriented toward designing an effective teaching package on racism and social inequality.

Monitoring the Effects of Historic Memory: What we set out to do

As CIRMA contemplated the content and structure of the campaign, IDRC`s Evaluation Unit offered to support CIRMA in conceptualizing and putting into place a comprehensive system for monitoring and evaluating the results, including the effects of the campaign. Monitoring peoples' reactions to the campaign was seen as a critical first step in assessing how the campaign might contribute to changing peoples' attitudes and behaviors towards racism. Through a series of peer support meetings and training in a planning, monitoring and evaluation methodology known as Outcome Mapping, IDRC worked with CIRMA staff in order to help them identify key boundary partners, those individuals and groups who the campaign most hoped to influence in present day Guatemala: high school teachers in public and private schools, regional university authorities, university professors in the regions and social and political leaders in NGOs.[22]

IDRC support assisted CIRMA to put together a global monitoring and evaluation strategy that included a series of monitoring and data collection instruments drawn from Outcome Mapping and other methodologies (questionnaires for dialogue participants, records of attendance at the expo, weekly reports by expo museum guides, interviews of expo attendees and monitoring national and international press coverage of the campaign). All of this data was collected and tracked through specially designed software. Monitoring efforts were largely focused upon the

---

[22] International Development Research Centre (IDRC). *Annual Corporate Evaluation Report*, Evaluation Unit, International Development Research Centre, 2008.

`*Why are we the way we are*?` museum expo which was considered by CIRMA to constitute the heart of the national campaign operation. IDRC and CIRMA each had interests – some of them shared - on issues of learning and accountability for results around the museum expo. CIRMA wanted to use its research to catalyze a process of national reflection and deep social transformation. IDRC wanted to know if its funding was achieving its desired outcomes and was also interested to see how Outcome Mapping might be successfully grafted onto a process for transitional justice and reconciliation. We also harbored the hope that data harvested from the monitoring and evaluation framework could be used to inform CIRMA`s continued research on inter-ethnic relations and social reconciliation in Guatemala.

What actually happened?

Results from the first two years of the deployment of the museum expo in three regions of the country exceeded expectations, generating a wave of reactions both inside Guatemala and internationally. Between mid-2004 and mid-2006, the expo reached more than 117,000 visitors (nearly 1% of the national population). More people visited the exposition than visited Guatemala's other 22 museums combined over the same period of time. The campaign received support from seventeen national and international donors and from of a broad cross-section of Guatemalan society, an extremely rare achievement in this post-conflict society.

The data gathered by the campaign indicated that the effort was received in general with high levels of acceptance and legitimacy. About 90% of the visitors to the exposition responded "positively" while 5% on either side viewed it as either too "light" or too "radical."[23]   Reactions tended to be positive although a minority of the visitors were defensive and negative. The range of reactions to the expo underscore the deep complexity inherent in recasting historic memory and cultural identity, as illustrated in the following descriptions:

---

[23] Adams, *op cit* at 5.

Reflections on personal identity and existing ambiguity:[24]

> "After visiting the expo, I thought more about it, and I am in fact, indigenous because my grandmother came from…"

Validation and feelings of comfort with the expo`s message and content:

> "At last, somebody is telling it like it is."
> "I was a solider in the 1980`s and I had to leave Guatemala, and that made me think about our living situation."

Surprise, doubt and curiosity around the issue being presented in a new light:

> "I was foolish to teach absurd things to my students; the true history is quite different, and here I finally recognized that."

Discomfort and a desire for more profound and convincing arguments:

> "We need to talk about structural discrimination in this country;"
> "They don`t say anything about reverse discrimination."

Fear in the face of the explicit presentation of the problem of racism:

> "This expo is a double-edged sword, because by trying to create awareness it can also awaken a sleeping consciousness that could turn into a monster."

Annoyance and denial:

> "What are you trying to do with this?"
> "I don`t think things are quite the way you show them."

The Campaign also received broad support from segments of the population that seldom agree on any issue. This could be seen in the makeup of the National Committee and Regional Support Committees, which included representatives of the central, regional, and municipal governments, academics, and leaders of diverse social organizations and private enterprise. Media support – written press, television, and radio – went well beyond simple reporting. In fact the media became key allies, almost always providing the campaign with ample and positive

---

[24] The following reflections are taken from Tani Adams, *Notes on the social impacts of the campaign to date, lessons learned and challenges for the future.*" Internal CIRMA document, 2007.

coverage. In 2004 alone, publicity firms donated more than $800,000 in free radio, TV and billboard ads in addition to other types of support such as the provision of bottled water.[25]

The exposition also awakened the interest of diverse international organizations working in the field of human rights and research and education in post-conflict societies. A number of international foundations requested presentations on the campaign or visited the expo with a view to documenting it as a model for reconciliation and educational innovation in a host of countries including Ethiopia, Eritrea, Serbia, Romania, Macedonia, Burma, Cambodia, Bolivia and Argentina.[26]

What was learned about historic memory and education in Guatemala?

In developing and implementing a monitoring and evaluation strategy for the museum expo, CIRMA and IDRC faced all of the challenges outlined in the first part of this paper, in addition to a number of other unexpected methodological and contextual hurdles. These were navigated with varying degrees of success. In some instances, CIRMA was able to collect useful data that yielded some learning around the effectiveness of the campaign and the nature of ethnic relations, as well as how to approach the issue more effectively and for greater reach. Four important pieces of learning stand out:

Mass exposure to the Exposition: The Campaign reached 25% of the local population and 40% of the school population in two of the three Guatemalan departments where the expo was set-up, creating an important unanticipated effect. By saturating a specific population, the exposition apparently created a neutral external reference point enabling people to safely discuss what had been a virtually taboo subject. This is an effect that the expo will seek to repeat by focusing on a critical mass of one social group during the next five years.[27]

---

[25] Adams, *Proposal, op cit* at…
[26] IDRC*, op cit* at ...
[27] Adams, *Proposal op cit* at 6.

Young people are a critical group and are especially open: The original idea of targeting the

expo to young people from 13-25 years of age was validated, as this group generally showed a

special ability to reflect about their ideas and reality. However, the expo organizers also observed

that children from ten years of age were just as responsive. Despite their openness, however,

these young Guatemalans would seem to require more formal facilitation to help them to

overcome a natural tendency to avoid this difficult issue.[28]

Teachers are critical and especially resistant: At the other extreme were the teachers,

who as a group were among the most ambivalent and resistant visitors to the expo, with a

tendency to stifle the reflection that their students were generating on their own. This is perhaps

not surprising; the structural discrimination that underpins Guatemalan society is sustained no

less by the education system than by the police or judiciary, the more common target institutions

of transitional justice.[29] With the re-launch of the expo in February 2009, the focus has been on

the development of a pedagogic strategy both to help teachers to digest the issue and to become

constructive participants in post-exposition dialogues and investigation in the classroom.[30] In

using the expo in a more directive, targeted manner, the expectation is that teaching professionals

in Guatemala can begin to move away from more traditional forms of emphasizing historic facts,

not enquiry, and that students will begin understand and problematize why accounts of the same

event might differ.[31]

Maximize the Exposition's efficiency and efficacy: Finally, an analysis of the costs

involved in the first phase of the deployment of the expo enabled CIRMA to develop a plan to

maximize financial and programmatic efficacy and efficiency for the final phase. An analysis of

---

[28] Adams, *op cit* at 6.
[29] Cole, *op cit* at 6.
[30] Adams, *op cit* at 6. Reference also to new the Pedagagic Strategy project proposal submittted to IDRC.
[31] Cole, *op cit* at 12.

reactions to different parts of the exposition also allowed CIRMA to identify certain changes to make to presentation of content and form of the installation.[32]

Despite the encouraging advances noted above, IDRC and CIRMA missed some major opportunities in efforts to track and understand the effects of the museum expo as a tool for social reconciliation. In mid-2007, CIRMA closed the expo in order to evaluate its results and plan for its re-design and re-launch in 2009. As part of that process of reflection, IDRC commissioned an external summative evaluation of the monitoring and evaluation framework experience. While many important lessons emerged from the evaluation, I will restrict myself to sharing three major learnings:

1.   The critical importance of defining use of findings

During the planning phase, much time was invested in working with CIRMA personnel in order to define how the data and analysis that would emerge from the monitoring and evaluation framework would be used and who exactly would use it. Based on the premise of Utilization Focused Evaluation,[33] the M & E framework endeavored not to dictate any particular evaluation model, method, theory or use of findings.[34] Rather, the expectation was that in guiding CIRMA through an interactive process, they would be better positioned to collect the data that the organisation would need for accountability purposes vis-à-vis its donors and for learning purposes around the effectiveness of the expo.

The implementation of the framework over a period of two years gave rise to a series of unexpected outcomes. The first, was a lack of agreement within CIRMA whether the framework and data collected should be used to inform decision-making for improving the operations of the expo thus informing program effectiveness and fulfilling accountability requirements with

---

[32] Adams, *op cit* at 6.

[33] Utilization-Focused Evaluation (UFE) starts with the premise that evaluations should be judged by their actual utility and use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything is done, from beginning to end, as this will affect use of findings. See generally, Michael Quinn Patton, *Utilization Focused Evaluation. The New Century Text*, edition 3, 2008.

[34] Micheal Quinn Patton, *Utilization Focused Evaluation Checklist*, 2002 at 1.

donors; or whether the data should be used to feed and inform CIRMA`s research activities. Although the data being yielded could arguably serve both purposes, the reality is that these are two fundamentally different groups of uses and users. This lack of clarity generated ongoing organisational tensions between academic researchers involved in the project and those who held very operational responsibilities.[35]

A second (and related) difficulty was the absence of spaces within CIRMA for analysis of and reflection about emerging data. The daily operation of the national campaign was an all consuming activity for CIRMA and its staff. The massive amounts of rich qualitative data that was collected on the immediate effects that the expo was having on visitors`attitudes tended to be inconclusive, due to weak baseline data. Despite this problem, informal daily exchanges between staff allowed CIRMA to incorporate some changes during execution and did significantly contribute to the re-design of content and process for the re-launch of the expo in 2009.

2.   Balancing power relationships

The challenges around use were exacerbated by the need to grapple with power differentials at two different levels: Externally, between CIRMA and the seventeen national and international donors who were financing the campaign; and internally between CIRMA staff who were directly involved with the Expo and those who continued to be occupied by CIRMA`s traditional research and organization activities.

Throughout the campaign, CIRMA was forced to spend an inordinate amount of time ensuring that its unwieldy vessel of donor relationships stayed on an even keel. Multiple donors with different reporting requirements, called upon CIRMA to report on diverse outcomes (and impacts) with differing data. This resulted in an ongoing tug of war between those who valued quantitative vs. qualitative data. IDRC unwittingly played a part in this drama: the summative evaluation of the project commissioned by IDRC in 2008 discovered that CIRMA staff involved

---

[35] CIRMA, Final technical report for project `Global Monitoring Proposal for CIRMA`s National Campaign for Dialogue and Interethnic Relations, 2008 at 23.

in the monitoring and evaluation work felt obligated to maintain an exclusive focus on Outcome Mapping tools – at a time when additional more quantitative tools were needed - because IDRC was funding the M & E project.[36]

In addition, part way through the project, the coordinator of the monitoring and evaluation work departed, taking with him his knowledge of Outcome Mapping. While the new coordinator made laudable efforts to get up to speed on the methodology, his research background was decidedly quantitative in nature. Although he managed to collect an impressive volume of quantitative (and largely demographic) data, opportunities to drill down to uncover qualitative considerations were lost. Nevertheless, much of the quantitative data collected did allow CIRMA to speak with authority to the outside world about the breadth and reach of the campaign.

Within CIRMA, both the national campaign and the monitoring and evaluation project generated unexpected tensions. The high profile of the campaign and related donor enchantment resulted in the creation of two camps groups within CIRMA: Those who were in the campaign and those who were not. The campaign was like a donor magnet, reproducing within CIRMA all of the worst deformations of international aid. The management of the daily of operations of the monitoring and evaluation project was highly centralized in the two coordinators who oversaw this work. Despite their best efforts, the coordinators were unsuccessful in generating excitement and building ownership for monitoring activities. The museum guides who were working in the expo collecting and entering data saw this work as a burden and imposition at the end of a long day. This was exacerbated by the fact that the data collection system was too large and ambitious. Data was being collected merely for the purpose of data collection in the hopes that in the future, it would serve the purposes of research.

3.   Theories of Change can be illusive when new ground is being broken

---

[36] Natalia Ortiz. Participatory Evaluation of the Monitoring and Evaluation System of the National Campaign for Dialogue and Interethnic Relations `Our Diversity is our Strength`, 2008.

Although both CIRMA and IDRC had a general idea of the expo project`s theory of change, we did not spend enough time actually articulating our assumptions and mapping out the potential complexities and variables – psychological, social/class, gender, political – that would influence the cognitive processing by individual visitors who viewed the museum expo. The implicit theory of change behind the exposition was probably something along these lines:

"Visitors to the museum expo will be exposed in a non-threatening and dynamic way to a comprehensive and historical accounting of the evolution of interethnic relations in Guatemala. Exposure will cause them to question their own construction of their identity as a `Guatemalan`. This questioning will catalyze behavior changes in their everyday activities and interactions in which considerations about race play an important role."

In hindsight, it is now easy to see that this theory of change was far too simplistic. Using the expo to induce behavior change among individuals was viewed as an important proxy for assessing social change and was one of the motivating factors behind the selection of Outcome Mapping as a central methodology. A more rigorous, participatory and externally validated process for thinking through the theory of change underpinning the expo would have allowed us to view and understand the expo as a tool for cognitive reframing (confronting individuals with information discrepant or contradictory to their expressed attitudes or self-image to induce cognitive dissonance and create opportunities for re-framing and re-organisation of these attitudes).[37] As it was, we missed the mark, and in this case, behavior change was never an appropriate indicator of social change – mainly because the viewing of the expo was in itself a passive activity which would have needed to be accompanied by clear strategies for social action if actual changes in behavior were to be the final outcome.

Poor mapping of the theory of change is a common trap into which we ought not to have fallen and IDRC`s accompaniment and mentoring of CIRMA could have been tighter. However, given the cutting edge nature of the national campaign and the fact that both CIRMA and IDRC were breaking new ground, it was only through the evaluation process that we were able to fully grasp the complexity of the theory of change underlying the expo and its use. This error has now

---

[37] Ilana Shapiro, *Theories of Change*, January 2005, p.3

been rectified. With its re-design and re-launch in February of 2009, the expo is now being used as a tool for training primary and secondary teachers and public servants on issues of race relations.

Concluding remarks

Pending.

Selected Bibliography

Tani Adams, 2007. *Proposal to Develop an Educational Strategy or Teaching about Racism and Ethnic Equality in Guatemala: Por Qué Estamos Como Estamos? 2007- 2011*, Internal document of the International Development Research Centre..

American Evaluation Association. *Program Evaluation Standards*. Accessed at: http://www.eval.org/EvaluationDocuments/progeval.html

Brahm, Eric. 2007. "Uncovering the Truth: Examining Truth Commission Success and Impact," *International Studies Perspectives* 8: 16-35.

Call, Charles. 2004. "Is Transitional Justice Really Just?," *Brown Journal of World Affairs*, 11(1): 101-113.

Centre for Global Development. 2006. *When will we ever learn? Improving lives through impact evaluation*, Report of the Evaluation Gap Working Group.

Duggan, Colleen. 2008. "Challenges and Prospects for Evaluating Transitional Justice". International Journal of Transitional Justice Board Conference. Capetown, May 28-30. Accessed at: http://www.idrc.ca/en/ev-135573-201-1-DO_TOPIC.html

International Development Research Centre-Center for the Study of Violence and Reconciliation. 2007."Evaluating Experiences in Transitional Justice and Reconciliation: Challenges and Opportunities for Advancing the Field," *Workshop Report*. Ottawa: CSVR-IDRC, April. http://www.idrc.ca/en/ev-114553-201-1-DO_TOPIC.html

Mendeloff, David. 2004. "Truth-Seeking, Truth-Telling, and Postconflict Peacebuilding: Curb the Enthusiasm?" *International Studies Review* 6(3): 355–380.

Neufeldt, Reina C. 2007. " `Frameworkers` and `Circlers` - Exploring Assumptions in Peace and Conflict Impact Assessment". Berghof Research Center for Constructive Conflict Management. August. http://www.berghoff-handbook.net, Accessed October 7, 2007.

Patton, Michael Quinn. 2008. *Utilization Focused Evaluation*.

Phuong Pham and Patrick Vinck. 2007. "Empirical Research and the Development and Assessment of Transitional Justice Mechanisms," International Journal of Transitional Justice, July (1): 231 - 248.

Ravallion, Martin. 2009. Should the Randomistas Rule*?*, *Economists Voice*, February. www.bepress.com/ev   Accessed April 15, 2009.

Shapiro, Ilana. 2005. *Theories of Change*, January. www.beyondintractability.org/essay/theories_of_change/ Accessed, April 23, 2009.

Smutylo, Terry. 2007 "Outcome Mapping: A method for tracking behavioural changes in development programs". *Institutional Learning and Change* (ILAC) Brief 7, August.

Thoms, Oskar N. T. et al. *"The Effects of Transitional Justice Mechanisms. A Summary of Empirical Research Findings and Implications for Analysts and Practitioners"*

United Nations Security Council. The rule of law and transitional justice in conflict and post-conflict societies. Report by the Secretary General. 23 August 2204. S/2004/616.

Vinjamuri, Leslie, & Jack Snyder. 2004. "Advocacy and Scholarship in the Study of International War Crimes Tribunals and Transitional Justice," *Annual Review of Political Science* 7(May): 345–362.