Measuring Security Progress

Politics, Challenges and Solutions

Summary report

Measuring Security Progress

author: **Megan Price**, Research Fellow Conflict Research Unit, Clingendael
date: January 2015

Table of contents

Introduction

Held 20 November 2014 at The Netherlands Permanent Representation to the UN[1]

Benchmarks, metrics, results, impact and evaluations have been trending topics within the international security community. Now, 2015 promises to witness a crescendo in this mounting wave of attention for data, as the Sustainable Development Goals are further specified into targets and indicators. In recognition of the salience of the topic and need for reflection, in late 2014 an expert meeting of policymakers, practitioners and researchers was convened to discuss the political nuances and operational challenges of designing, applying and interpreting security progress measurements. This report seeks to capture the debates, lessons and innovations that emerged from the day's discussions. The meeting, which was held under the Chatham House Rule, benefited from the deep and varied experience of the speakers, as well as the critical reflections offered by the participants. The organizers would also like to acknowledge the staff of The Netherlands Permanent Mission to the UN, for assisting with coordination and generously hosting the event in their New York mission.

> In 1907, a New York Times headline read: "Soul Has Weight, Physician Thinks".[2] Below this rather droll title ran a story summarizing the findings of Dr. Duncan MacDougall, who sought to determine the mass of a human soul by weighing people, literally, on their deathbed. After six observations, he reported an average loss of 21g upon his subjects' death. Unable to physically account for the deficit, the doctor concluded this was the mass displaced by the soul leaving the body. His search for alternative explanations, we could guess, may not have been thorough.

This brief anecdote works to illustrate that, historically, the abstract nature of a concept has not deterred attempts to capture it in metrics. Moreover, empirical data is often interpreted through the lens of a dominant narrative and can incline, sometimes steeply, towards confirmation bias.

While many may smirk at the above story of antiquated pseudo-science, how are current efforts to measure 'security progress' or 'rule of law' regarded? How sound are the correlations we currently draw between measurable observations and these intangible concepts? If we tally up an increase in police recruitment, and note a concurrent drop in reported muggings, do we readily conclude that 'more police equals more security'? How thorough are our own searches for alternative explanations? What does our choice of indicator tell us about what we *expect* to find?

These were a few of the queries that brought together security practitioners, researchers, policy officers, criminologists, donors, and UN representatives to jointly discuss state-of-the-

art dilemmas and innovative strategies for defining and qualifying security progress. The meeting, held in New York, was divided into two sessions. Session I was spent looking at how security indicators reflect specific priorities and are prone to reinforce a particular security narrative. Session II then turned to look at innovative strategies for compiling and conveying security information in challenging data environments. The structure of this report mimics the day's program, [3] providing a synopsis of each discussion and sketching the contours of an emerging topic of debate.

BOX 1: TAKEAWAY POINTS: HIGHLIGHTS FROM THE DAY'S DISCUSSION

- Negotiating what evidence 'counts' is part and parcel of determining who is empowered to influence the narrative of 'progress', and who is granted the authority – and the resources – to respond. Therefore, indicators have profound political implications.

- Indicators can create incentives for particular behaviors and empower certain actors within a sector; they not only reflect change, but also influence its course.
- Many programs measure a security or justice context in terms of what it 'ought to look like', while missing the reality that people actually experience. Current tools for monitoring and evaluation have generally failed to address this normative tendency.

- Sustainable security and justice is a matter of political progress, yet our current tracking methods focus on technical gains made. Emerging research offers some ways forward.

- Data is more likely to stimulate positive reform when it exposes issues in a way that can be managed, and that is presented in a constructive rather than a judgmental tone.

- In some cases, the amount of trust people have in certain information is more relevant than whether the data is from an 'official source' or even whether it is accurate.

- There is an inherent tension between endorsing global norms and standards, and simultaneously acknowledging the value of locally derived metrics and nationally driven processes of change. Though both are valid and useful, they can operate in competition.

Currency of the topic: All this talk of measurement

The international community is investing attention and resources into developing tools to better account for the impact their support has had on security in host countries (trailing by a

few years similar trends in development assistance).[4] Donor efforts have been concentrated into monitoring and evaluation (M&E) mechanisms and 'upstream' accountability frameworks. Global indices of peace, fragility, et cetera, have proliferated, with the aim of depicting periodic shifts or stasis in countries' relative levels of violence or stability. These trends collectively point to a growing international consensus on the need to observe and record changes in 'security'.

The consensus, however, appears to end there. Although many actors are eager to illustrate and track security progress, there is scant agreement on what 'security' is or what signifies 'progress'. As one participant pointed out, the term 'security' was itself too contentious to be included in the Post-2015 Sustainable Development Goals, and was replaced with the more politically palatable 'stable and peaceful societies' goal. Though equally difficult to measure, 'stability and peace' engender less cynicism over sovereignty issues, helping to increase buy-in from a broader range of actors.

However, as the day's discussions revealed, this is not a problem of semantics that dialogue could attempt to resolve. Rather, it is a reminder that developing indicators of security is not unlike attempting to *define* security. Consequently, the goal is not to broker universal agreement on the 'correct' metrics of security; it is to build awareness that measuring security has profound political implications. The day's discussions provided a wealth of examples, as elaborated hereunder.

Session I
What 'counts' depends on who's counting

To untangle this layered topic, it is helpful to first distinguish the different levels at which indicators are defined, and the various audiences and purposes they serve. A global benchmarking specialist requires a different data set from the one that a police chief in Port-au-Prince or a UNDP programming officer in New York requires. When looking at a particular stock of data, it is important to consider who is gathering this 'evidence', and how they intend the results to be used. This serves to pull out distinct threads of the conversation, and avoid confusing means with ends. That is, data is not collected for data-sake; it is collected to help tell a story. Negotiating what evidence 'counts' is part and parcel of determining who is empowered to influence the narrative of 'progress'.

In the morning session, speakers discussed indicators endorsed in international advocacy forums, metrics designed for country-level monitoring, and information disseminated through public media outlets in conflict-affected settings. As outlined in the subsequent sections, the ensuing conversations delved into important debates particular to one level or another, but also revealed parallel observations and similarities occurring across all three arenas.

Framing a global agenda

For about 15 years the global community working to reduce violent deaths has used annual homicide rates as the main proxy indicator of violence. In 2011 alone, more than 526,000 violent deaths were estimated across the globe.[5] The blunt figures have been effective in galvanizing an international agenda for action, and establishing targets for reducing the toll. However, the comprehensive totals were repudiated by governments that felt such 'score-keeping' was an unfair and simplistic reflection of their complex domestic situations. In short, there was need for nuance.

As the international community diversified its analyses of violence, emerging schools of thought advanced more sophisticated ways of framing the issue. For example, those assessing global insecurity primarily as an outcome of armed conflict focus their calculus on annual battle-related deaths. Proponents of armed violence reduction count up a broader swathe of violent deaths, including those resulting from criminal violence, demonstrating that non-conflict fatalities represent a deadlier scourge than casualties of war. The public health community, meanwhile, seeks to track all violent mortalities, including accidental homicides, suicides and deaths at the hands of law enforcement. It no longer suffices to know how many

people were killed; different communities of practice are demanding higher-resolution data, providing details on who, when, how, and where.
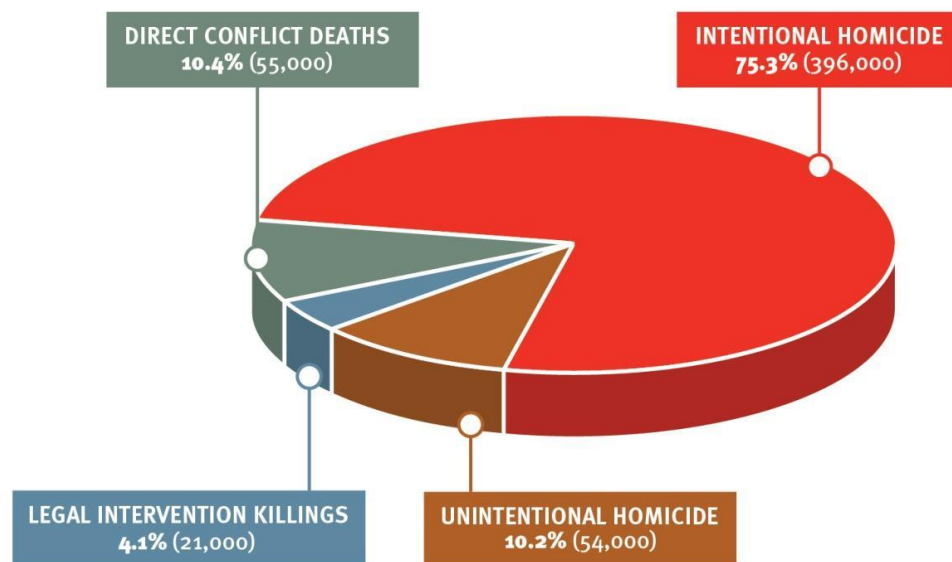

Figure 1. Modified from Geneva Declaration, Global Burden of Armed Violence, 2011, p70

Quickly, the international dialogue grew more cacophonous, as each new frame brought new lexicons and processes. While this has led to what one participant termed "a global burden of indicators", disaggregating sources of lethal violence, or groups disproportionately affected by it, can help to calibrate more effective responses. This can enable different communities of practice to engage their unique expertise and approaches. Defining new metrics at the global agenda-setting level provides experts with the unifying language necessary for mobilizing international action. In a sense, the propagation of various indices contributes to getting all hands on deck.

However, this coin has two sides. Collecting a certain set of data is not simply about articulating the problem; it is also instrumental in framing the appropriate course of action. Depending on the data collected, violence can be compellingly presented as a corollary of, inter alia, small weapons proliferation, gang culture, political repression, economically motivated crime, or misogyny. This has implications for which communities of practice will be expected – *and supported* – to respond (and which will not). Setting global standards for reporting can reinforce certain countermeasures as the norm, perhaps to the exclusion of more locally generated responses. Interestingly, similar observations were echoed at the programming level as well.

At the 'country level'
Several speakers reflected on how selecting indicators often meant navigating the competing narratives and priorities of different stakeholders. The most obvious dialectic in this regard is

that of the host government and the donor. Here, defining program outcome indicators can run up against divergent expectations of the scope and speed of change, contrasting emphases on state or 'people-centered' security, or differing views on the means necessary to pursue program goals.

Such political–strategic dilemmas can unexpectedly arise during the creation of a program monitoring framework – an ostensibly technical exercise. In order to reduce attribution error, expected signs of progress are tightly knit with how and where resources are to be channeled. Thus a program outcome indicator can become a lightning rod for latent differences in strategic vision. One speaker discussed an experience in Ethiopia where the donor's plan to record program outcomes at the village level exposed a cleavage with the government's desire for nationally visible results, reinforcing central security capabilities rather than those of local agents.

This example underscores another important observation that emerged from the presentations: defining indicators can also become embroiled in domestic competition over bureaucratic culture and sector leadership. Using domestic administrative data to track progress can, deliberately or inadvertently, strengthen one national actor's position vis-à-vis that of others. For example, gathering national sector-level data can help build a central ministry's statistical capacity and its authority to mandate and manage comparable data collected by various field offices. On the other hand, working with local-level indicators can help empower decentralized agencies to take the lead in defining and collecting information they deem relevant for progress on their terms. Either method for collecting data is apt to bolster or disrupt bureaucratic power balances within a sector, though external players may not be sensitive to these dynamics. However, negotiating the right balance among these actors' inputs may help encourage the uptake of data throughout an administration.

In addition to vertical relationships, indicators can also influence horizontal politics among different agencies. In practice, police recruitment numbers are commonly used as a way to benchmark progress in the rule of law sector. However, this risks an asymmetric focus on the sharp end of enforcement, to the detriment of other custodians of justice, such as public defense attorneys, or human rights inspectors.[6] As one presenter described it, this can "create winners and losers" within a sector. Beyond bruised egos, such benchmarks can have significant implications for where the international community and the domestic government are encouraged to invest their resources.[7]

The crux of this discussion is the fact that 'what gets measured gets done'. Indicators create incentives for particular behaviors or activities; in essence, shaping the environment they are meant to describe. Here, it is crucial for policymakers and program officers to be aware of the norms and behaviors they are endorsing through the data they collect.[8] What is more, consideration must be given to what data may be missed by deferring to international reporting standards, rather than locally derived metrics.

Taking this idea one step further, it is worth considering how data collected by local administrations may provide a wealth of insight into the interests, priorities and ambitions of local security and rule of law agencies. When dealt with mindfully, such data can provide

entry points to discussing reform. One presenter shared his experience using data collected by Jamaican police[9] to initiate more strategic conversations about managing performance priorities and notions of accountability. The potential here for reform-oriented capacity-building is often overlooked when data is seen as a technocratic obligation, and not recognized for its strategic, management uses.[10]

Communicating to the public

An area in which data is often recognized for its strategic use, or narrative role, is that of the media and public discourse. Getting security-oriented messages out to communities and local populations is a challenge even in comparatively stable societies. The implications of inflammatory information or misinformation are potentially destabilizing, particularly in volatile settings.

One presenter offered some key insights into the complexities of how such information is presented in countries that host UN peacekeeping missions. Most UN Status of Forces Agreements, or SOFAs, contain a provision for a UN public broadcasting unit. These units, if successful, can play an important role in contributing to the narrative of the country's security progress, helping to shape public perceptions. However, that success relies to a large extent on the reputation the broadcast is able to build among the national population. According to the presenter, this is first and foremost determined by the ability to convey messages in a local tone, not only broadcasting in the local language(s), but also employing local journalists and presenters, and using the media conduits most readily available to communities across the territory, often radio.

Power struggles can and have emerged between host governments and the international community when determining who owns the content of public broadcasting. Considering the investments and reputations at stake, both actors have a lot riding on at least the general impression of progress, though their narratives may diverge on the reasons for gains made. However, ensuring that information provided through public media is not only accurate but trusted requires, as the presenter phrased it, the willingness to deliver 'bad news' in a straightforward manner. Moreover, trust can also be frayed by association with external actors, namely the UN. In some cases, this has required the UN to create a perceived distance between itself and its broadcasting system in order to help preserve the legitimacy of its radio programs, underscoring the complex relationship between information and politics.

Closing reflections: politics of measuring, and the measuring of politics

The very act of measurement is a politically informed choice. The selection of what is worth documenting and deciding how to present this data are guided by expectations of how the information should be received, what assumptions it should confirm and what actions it should instigate. In this way, it is important to consider how indicators not only reflect change, but also work to influence its course.

If we accept this premise, what then are the implications of measurements that remain primarily focused on technical outputs and outcomes? Do they not only reflect but also

actively reinforce a technical approach to security and rule of law reform? Is it possible to reconcile the political nature of reform with the technical tools used to describe its progress? One speaker cited how current measurements of rule of law tend to frame the lack of security or justice as 'flaws' in the system, glitches that can be corrected with capacity-building or improved infrastructure. However, there is increasing acknowledgement among researchers and practitioners alike that insecurity and injustice are often part of the system, embedded in the resident design of power. Thus, the durability of any 'gains made' – for example, measureable reductions in violence, or increased access to justice – fundamentally requires political structures to guide and reinforce technical improvements. Unfortunately, the hardware is much easier to measure than the software.

The presenter compared tracking changes in the rule of law to charting the path of a sailing boat, moving with undulating currents, capitalizing on headwinds, occasionally treading water or being carried backwards. Meanwhile, most M&E frameworks, often developed with donor 'return on investments' in mind, seem to envision change as though moving along a railway track, progressing in linear fashion, generally predictably, perhaps even 'on time'. Political change, though, cannot be so trimly scheduled.

This brought up a second point: when to measure? Consideration of the time required for social transformation is a critical element in measuring changes in the rule of law, security or justice. Intended or unintended impacts of certain programs may not reveal themselves for decades, only coming clear in retrospect and with the benefit of historical analysis. And even within the lifespan of a program, expectations and goals are liable to shift with political winds, rendering irrelevant the baselines or indicators put forward in the incipient phases of the work. In sum, rigid short-horizon indicators deployed for periodic donor monitoring are, by and large, ill-suited for measuring the slow, iterative nature of political change, which is better captured by analytical description than by quantitative inventory.

With these points in mind, it was proposed that new approaches to measurement should include a more nuanced understanding of change occurring in gradual, irregular phases. Rather than tracking singular indicators, baskets of interrelated indicators could be used to analyze broader patterns of change over time, and to sketch the boundaries of what is possible in a given period and context.11 Similarly, gaining a better appreciation of the conditions that may spark or enable a political shift, or 'critical juncture',12 can be as important as describing its initial outcomes. Understanding the dynamics that led to and sustained the US Civil Rights Movement is more informative than an accurate tally of the number of schools integrated between 1962 –'63. Developing such approaches to measurement will require a stronger reliance on qualitative tools such as scenario planning and political economy analysis to better represent – or make visible – the complexity of change.

Session II
Finding innovative ways forward

The afternoon session was launched, nimbly picking up on the topic of 'visibility' and its implications for the ability to monitor and measure. In a stylized system, information on inputs, outputs and outcomes is evident, enabling informed analysis of ultimate impacts and how these might be adjusted. Outside the lab, by contrast, what we can actually 'see' is much more limited. This is particularly true in fragile and conflict-affected settings, where inaccessible terrain, weak national structures, and unforeseen dynamics create challenging data environments.

Most of the time, when working on the provision of security or the advancement of the rule of law, we have much less information than we need to be able to draw solid conclusions. In spite of this knowledge gap, or perhaps because of it, those seeking to gauge progress often measure the security or justice context in terms of what it 'ought to look like', while missing the reality that people actually experience. For example, country justice assessments, dutifully following global standards, often report on the independence of judges, the number of courthouses, or volume of backlogged cases, but may fail to recognize or register the value people place on local forms of traditional arbitration. To address this chronic blind spot, innovative feedback loops are needed to reveal the existing systems that determine or influence people's experience of security and justice, and to insert this information back into the objectives and design of reform programs.

This session delved into the challenges of, and innovative approaches to, detecting intervening variables, factors of difference, and the incalculable inputs that shape local realities. With such tools, indicators could be better designed to speak to local systems of security and justice, rather than measuring 'gaps' that distinguish current realities from imagined futures. These discussions covered the design of public perception surveys, tracking changes in informal justice provision, and the use of mobile technology in conflict and post-conflict situations.

Perception surveys: "Don't measure what can't be managed"

Population-based perception surveys have been embraced by reform practitioners as a methodological miracle: a practical, quantitative tool to gauge intangible, qualitative realities. Looking beyond technical outputs, they offer insight into the day-to-day experience of security and justice. Yet, despite their potential to start a conversation about the content and direction of reform, in practice, survey results are often used to hone a sharp critique or declare 'conclusive evidence' of a system's inadequacy. Wielded as a tool to contest power,[13]

surveys can goad antipathies among local authorities and entrench a reluctance to engage with the results, even among institutions relatively open to reform, effectively ending a conversation before it has begun.

The tendency for results to be used to embarrass or criticize is not inherent in surveys, but can be attributed in part to the tool's methodological limits. Working more like a weather vane than a compass, a perception survey can indicate the fluctuating moods of a population, but is less useful in orienting its user towards the 'right' direction. That is, the tool points more quickly to problems than to solutions. This creates dilemmas when the indicators designed to diagnose insecurity or injustice are more sophisticated or broader reaching than the capacities of the agencies designated to respond.

For these reasons, perception surveys are most expedient when complemented by focus group discussions, political analysis, and constructive engagement with local security and justice actors. On this latter point, it is remarkable that surveys rarely engage the actors whose conduct is to be influenced, namely police officers, judges, among others. Understanding such actors' incentives, interests and world views, and how these relate to those of citizens, can provide valuable insights into the viable space for reform. Or it may, in fact, help to leverage greater space for reform.

Illustrating this point, the presenter related an anecdote in which national police took up the distribution of a security perception survey themselves. While many participants (your author included) initially baulked at the idea, the presenter celebrated the occasion for police officers to actively engage their communities, asking them directly about their security perceptions and needs. The benefits of such a practice, it was argued, could prove more valuable than the pretense of objectivity. This underscored a key point. The goal here was not measurement itself, but rather to encourage national agencies to consider using local security perceptions to inform their policies and practices.

Such data is more likely to stimulate positive change when it can generate information precise enough to help local agencies decide when and where to deploy resources.[14] As an example of how to expose issues in manageable terms, the survey results below not only focus on a specific situation, but also highlight which particular neighborhoods have not benefited from the general trend of improvement. Though this data alone cannot explain why this is the case or suggest countermeasures, it serves to capture the issue in manageable terms.
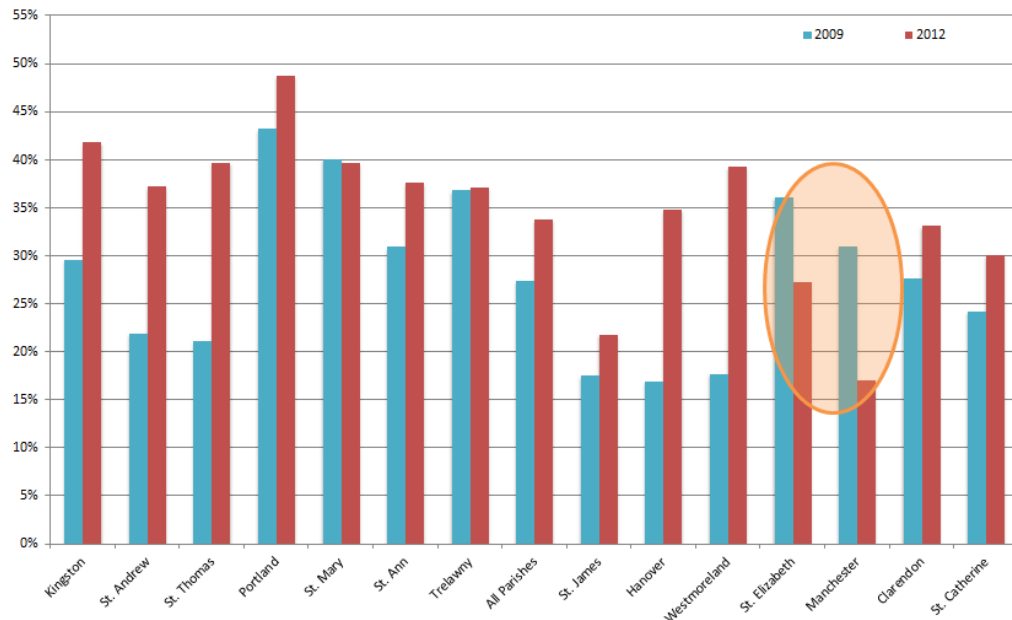
Figure 2. Modified from Foglesong, Better Servants of Development, 2014, p9

Another way to increase the likelihood that survey results contribute to reform is to be sensitive to the differences between perceptions, which can be volatile, and experiences, which are more concrete. Gathering information on people's actual experience of crime, violence or corruption, and the frequency of certain incidents, can help provide a more nuanced picture of what may be shaping people's ideas of security or the rule of law. Conversely, surveys hastily conducted may simply reproduce or give credence to misinformation and stigmas about particular agencies, groups or neighborhoods. This prompted one participant to remark, "In most cases, no perception survey is better than a poorly managed perception survey."

Lastly, surveys that facilitate in-country comparisons, either comparing localities, or measuring changes over years, are often better received by domestic agencies than surveys that invite comparisons with international standards or other countries. On the other hand, while international comparisons may give national actors less practical advice about how to manage their local contexts, global standards can stimulate useful national discussions and help to focus priorities.

Measuring non-state justice: formalization over function?

It is fairly easy to recognize the normative lilt in measuring security and justice reform against international (or perhaps more accurately, 'Western') standards of practice. While this is still compatible with what donors are willing to support, one speaker discussed how indicators that work from this narrow lens can actually prove counter-productive, specifically in terms of improving justice outcomes for certain populations.

For roughly the past 15 years,[15] the international community working on rule of law and justice reform has been coming to terms with the prevalence of 'informal, customary, hybrid and non-state' justice provision. These various systems are commonly assessed according to their alignment with the 'formal' state system, typically applying indicators such as: transparent record-keeping; referral or reporting to state law enforcement; and adherence to international human rights standards.[16] While such aspirations are certainly not specious, they may be somewhat misplaced when it comes to non-state justice systems, whose alignment with formal institutions is not necessarily intrinsic to their ability to deliver satisfactory justice outcomes.

Applying 'formal' justice criteria to non-state justice systems is part of a narrative that promotes formalization, typically as part of a larger state-building campaign. However, indicators that measure formalization as a proxy for justice progress are liable to assert state-building objectives over the aim of improving justice outcomes for people. Formalization indicators, such as those above, rarely provide insights into the nature of people's justice needs, much less how those needs may best be served by either formal or non-state systems. Moreover, the way in which progress is measured does not always account for the negative outcomes of formalization, which in practice can often mean a reduction in people's options to pursue justice.

Some of these consequences are the result of taking the proposed solution, formalization, as the starting point. This approach precludes conversations about what is necessary to work towards improved justice outcomes, irrespective of the provider. However, the presenter proposed that working from a problem-oriented perspective could offer new pathways forward. This implies, first and foremost, empirical research into which specific justice problems, or sources of injustice, are to be addressed. Doing so requires combining both qualitative and quantitative analyses, and is not restricted to looking at the 'justice sector' per se. The second step involves looking at the constellation of authorities and agents designated with the power to manage or mediate the problem identified. Again, the approach does not presume who these actors or systems are (or should be), but rather opens the aperture of the lens for capturing a broader range of existing local structures and processes. From here, attention can be given to the need and scope for change, what structures maintain the status quo, how reform may be productively pursued, and on what terms it can be monitored. Once more it was stated that these questions must be iterated throughout a process of reform to ensure that monitoring is in step with ever-evolving reality.

Technology and information ecosystems: mobile 'truths'

The last presenter of the day reinforced the notion that within local systems, 'formality' may matter less than is assumed. With regard to local information ecosystems it may matter less than ever.[17] Modern information technology has rapidly expanded and decentralized the options available for receiving, disseminating and archiving security data, particularly in fragile or conflict-affected contexts. The presenter explained how this expansion necessitates "forward-looking models to comprehend local information systems" and their multifaceted impact on security dynamics.

Discussions of mobile technologies and security information typically gravitate towards innovations that amass critical data either in 'real time' or from remote locations, reducing the barriers to extracting information from traditionally challenging data environments. This is but half the story. What is less often discussed, though equally relevant, is how this technological leap forward has also reduced barriers for information to penetrate and circulate through these environments.

Understanding local communities' perceptions of security is inextricably tied to how information about threats or violent events are spread through local communication systems. Such information is increasingly likely to flow through 'trusted circles' such as Facebook groups, Bluetooth, or WhatsApp messages. In these closed systems of information exchange, confirmation bias and peer-to-peer trust diminish the demand for verification, and can increase polarization. In such circles, those with the strongest influence over the conversation are not always those with the most reliable information. Nonetheless, this information influences people's security perceptions and, more importantly, the decisions they make regarding personal safety.[18] A widely circulated rumor, even if unfounded, can spark a mass displacement or a violent mobilization just as quickly as a veritable threat. With this in mind, it is interesting to consider how the information available to and within communities can contribute to early warning systems, regardless of its accuracy.

Information flowing through these local communication ecosystems may not be reliable enough to predict or identify security trends or events. However, being aware of and able to access the messages being dispersed through these local systems can provide insight into community members' perception of a conflict, the significance they assign to (alleged) events, and the views they hold of those involved. These insights can greatly serve analyses of conflict drivers, inter-group tensions, and the potential fight or flight responses to rumors or incidents.

In this way, the presenter pointed to an overlooked dimension of the relationship between security data and mobile technology, demonstrating how information ecosystems offer a new paradigm for analyzing local security. The ability to predict disasters or violence, and to provide appropriate and timely support, is contingent on reliable information.[19] However, being able to predict and respond to communities' reactions to *perceived* threats requires insight into the information they are receiving and sharing. Here, veracity may be wholly beside the point.

Reflections going forward

When the topic of monitoring security and justice arises, a great deal of attention is currently focused on collecting the 'right' data. What metrics are most revealing? What data is most reliable? What are the most rigorous methods to gather and verify information in remote or even hostile settings? While these are pertinent questions, the meeting made clear that they should not delimit the boundaries of the discussion. What is more, these questions may be obscuring a crucial point. Underlying this preoccupation with technical precision is a presumption that conclusions drawn from the 'correct' data will be self-evident. Since, supposedly, 'numbers don't lie', measurements are often assumed to be somehow objective-by-association. This is a fallacy prevalent in M&E.

This simplistic view severely undervalues the critical role that analysis and interpretation play in designating which measurements are significant and what they signify.[20] But what is more, it ignores the role that power plays in determining who has the authority to assign meaning to fact.

When describing rule of law and security, the idea of an objective indicator is a myth. Measurements are tools intended for a purpose, and they inevitably reflect and promote a set of norms, expectations and assumptions. This instrumental relation between empirical data and normative affirmation is often overlooked or poorly articulated. The aim of the meeting, and the discussions captured in this report, has been to expose this relationship on multiple levels.

The material presented here is intended to provoke and contribute to further reflection and debate among the international community in the coming year, this International Year of Evaluation. Such debate is particularly relevant to discussions around the Post-2015 Sustainable Development Goals (SDG) indicators, which will set the future terms for global security and rule of law aspirations. As one speaker rightly pointed out, the very inclusion of topics like peace and violence reduction among the SDGs represents a significant step that was unthinkable a mere decade ago. But with the opportunities this policy achievement affords come unaddressed tensions sure to challenge policymakers who have lobbied strongly for including the 'stable and peaceful societies' goal.

The thrust of this challenge has been echoed throughout the report: the difficulty of reconciling local visions, ambitions and interpretations of security with more broadly germane international standards. Experiences from the International Dialogue on Peacebuilding and Statebuilding[21] provide useful insight into the inherent trade-offs and perplexing considerations involved in accommodating locally driven processes of change and bolstering

domestic leadership, while also endorsing global indicators to mobilize international resources around issues of shared relevance.

As stated early in the report, both internationally agreed standards and locally determined metrics are necessary; but they can operate in tension, or even competition. It can be analytically expedient to separate out the different purposes and audiences that indicators serve. (It certainly made for a tidier meeting program.) But these debates and discussions often prove much more difficult to extricate in practice. In this way, it is important to be ever-aware of what norms and expectations are embedded in data presented. As one speaker put it, empirical claims will not resolve what are – deep down – normative disputes. But they do have a way of steering the conversation, emphasizing certain questions and concerns while sweeping others aside. This should be considered the next time one reviews a police–population ratio or stability index that focuses strictly on violence but ignores corruption. The question bears repeating. What does our selection of indicators and measurements tell us about what we expect to find?

Literature

[1] Organized under the auspices of the Knowledge Platform Security & Rule of Law, this meeting was part of the 2014 Event Series exploring frontier issues that currently confront international efforts to strengthen security and rule of law. Further information can be found on the Platform's website: www.kpsrl.org

[2] *New York Times*, 11 March 1907. http://query.nytimes.com/mem/archive-free/pdf?res=9D07E5DC123EE033A25752C1A9659C946697D6CF

[3] The program, organized under the Chatham House Rule, can be found on the Knowledge Platform Security & Rule of Law website: http://www.kpsrl.org/

[4] Incidentally, 2015 was officially declared as the International Year of Evaluation (EvalYear) at the 3rd International Conference on National Evaluation Capacities held in São Paulo, Brazil, 29 September to 2 October 2013. The aim of EvalYear is to strengthen the demand for and use of evaluation to inform public policies. http://www.mymande.org/evalyear

[5] Geneva Declaration, *Global Burden of Armed Violence: Lethal Encounters,* Cambridge University Press, 2011.

[6] Consider, for example, that UN mission benchmarks are often pegged to a target number of national police officers, rather than the human rights records held by those security forces.

[7] See also, E. McCandless, "Wicked Problems in Peacebuilding and Statebuilding: Making Progress in Measuring Progress Through the New Deal", *Global Governance*, v.19 no.2, 2013, pp227-248.

[8] This point is often used to argue against indicators such as 'number of arrests' or 'number of convictions' to monitor the effectiveness of rule of law actors, given that such indicators can incentivize arbitrary incarcerations.

[9] An overview of the project can be found on the website of the Harvard Kennedy School Program in Criminal Justice Policy and Management. www.hks.harvard.edu/programs/criminaljustice/research-publications/measuring-the-performance-of-criminal-justice-systems/indicators-in-development-safety-and-justice/countries-and-partners/jamaica

[10] As one participant pointed out, UN peacekeeping operations lack a specific budget line for monitoring security progress, and tracking 'gains made' is often an addendum to project officers' regular workload. In the absence of specialized analysis, monitoring thus tends to focus countable outputs. For more on the need to build monitoring capacity within peacekeeping and peacebuilding organizations, see S. Stave, *Measuring peacebuilding: challenges, tools, actions*, Oslo: NOREF, 2011.

[11] This short synopsis draws on the presenter's more in-depth research on monitoring the underlying processes of social transformation and rule of law, to be published in a forthcoming paper. (Meeting held under Chatham House Rule.)

[12] 'Critical juncture', associated with path dependency theory, in this case aligns with Robinson and Acemoglu's term to describe "major events that disrupt the existing political and economic balance in one or many societies". in *Why Nations Fail: The Origins of Power, Prosperity and Poverty*, New York: Crown, 2012, p431.

[13] Survey results have the potential to create political disruptions not unlike elections. Done on a large scale, or at strategic moments, surveys can appear as unofficial popular referendums on the performance of the police, military or justice institutions. This strategy often has clear designs on driving the security narrative.

[14] For a more comprehensive treatment of this topic, see T. Foglesong, *Better Servants of Development: Improving Surveys as Sources of Indicators for Public Safety*, Boston: Harvard Kennedy School, 2014.

[15] The presenter referred to the UN Transitional Administration Mission in East Timor (UNTAET) which was one of the earliest cases of international actors working with 'non-state justice systems'.

[16] One of the more nuanced descriptions of indicators for non-state justice systems is put forward by the VERA institute, though it ultimately defers to measuring in terms of formalization. VERA Institute, *Measuring Progress toward Safety and Justice: A Global Guide to the Design of Performance Indicators across the Justice Sector*, New York, 2003. pp65-73. http://www.altus.org/pdf/mptsj_en.pdf

[17] In this context 'formality' is meant to imply regulated public communication, such as mainstream news media or official public service announcements. In contrast, the presenter defined information ecosystems as "complex adaptive systems that include information infrastructure, tools, media, producers, consumers, curators, and sharers".

[18] A more in-depth discussion of this point can be found in A. Ayala, "The Dichotomy of Technology in Conflict: Beauty and the Beast" in *Communications Technology and Humanitarian Delivery Challenges and Opportunities for Security Risk Management*, European Interagency Security Forum, 2014.

[19] While accessing data has become somewhat easier, triangulating this information remains a key challenge – though it can be said that information technology has proved to be an asset for both purposes.

[20] The prevalence of this misperception may go some way towards explaining the dearth of analysis that some meeting participants described in international peace organizations' current monitoring frameworks and practices.

[21] See International Dialogue, *Peacebuilding and Statebuilding Indicators – Progress, interim list and next steps (for discussion*, 2013. Available online http://www.pbsbdialogue.org/documentupload/03%20PSG%20Indicators%20EN.pdf

Knowledge
Platform
Security &
Rule of Law