NEW ISSUES IN REFUGEE RESEARCH

Research Paper No. 214

Indications of progress? Assessing the use of indicators in UNHCR operations

Emma Dunlop

emma.dunlop@law.nyu.edu

July 2011

UNHCR
The UN Refugee Agency

Policy Development and Evaluation Service

Policy Development and Evaluation Service
United Nations High Commissioner for Refugees
P.O. Box 2500, 1211 Geneva 2
Switzerland

E-mail: hqpd00@unhcr.org
Web Site: www.unhcr.org

ISSN 1020-7473

Introduction

In recent years the United Nations High Commissioner for Refugees (UNHCR) has placed increased reliance on the use of internally generated indicators as a tool of governance. In an effort to improve project planning the organization has become a prodigious producer and consumer of indicators that variously map refugee conditions on a country level, within camps, in urban programs, and in returnee areas.[1]

Indicators are considered to be a boon for policy-makers charged with balancing protection and operational needs on a global basis, in that they recast complex fact-scenario issues into quantitative data that is susceptible to ranking and comparative analysis.[2] However the increased emphasis on indicators as a tool for governance also raises concerns regarding the production, use and dissemination of data. This paper seeks to apply insights of recent scholarship in the indicators field to governance by indicators in UNHCR operations.[3] The use of indicators by other actors in international refugee law falls outside the scope of this paper.[4]

Practice within UNHCR is particularly ripe for analysis given the organization's growing reliance on indicators in internal decision-making and the breadth of its operations. The *Global Appeal 2011 Update* estimates that more than 36 million people worldwide fall within UNHCR's mandate as "persons of concern".[5] The organization works in 125 countries,[6] and has declared a "long-term vision of establishing and systematically using a set of agreed standards and indicators for planning and measuring the impact of its operations".[7]

---

* I am indebted to Eyal Benvenisti, Sabino Cassese, Ryan Goodman, Meg Satterthwaite, Kate Horner and Hugh Atkin for their comments on this paper.

[1] See UNHCR Division of Operational Services, *Practical Guide to the Systematic Use of Standards & Indicators in UNHCR Operations* (hereafter *'Practical Guide'*) (2nd ed) (February 2006).

[2] See e.g. M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights" (2008) *Institute for International Law and Justice ('IILJ') Working Paper 2008/12*, p. 2; K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*, pp. 4-5.

[3] The use of indicators as a tool of governance is gaining increasing scholarly attention. Representative publications include K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*; N K Dutta, "Accountability in the Generation of Governance Indicators" (2010) *IILJ Emerging Scholars Papers*, New York University School of Law; M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights" (2008) *IILJ Working Paper 2008/12;* A von Bogdandy & M Goldmann, "The Exercise of International Public Authority through National Policy Assessment: The OECD's Pisa Policy as a Paradigm for a New International Standard Instrument" (2008) 5 *International Organizations Law Review* 241; K Davis & M B Kruse, "Taking the Measure of the Law: The Case of the *Doing Business* Project" (2007) 32(4) *Law & Social Inquiry* 1095.

[4] Indicator use beyond UNHCR includes the possible application of indicators as "Country Information" in domestic courts charged with making refugee determinations; indicators adopted by national governments, such as the "Indicators for Integration" proposed for the UK Home Office; and regionally-based indicators, such as MIPEX, which measures policies to integrate migrants across 25 EU and 3 non-EU countries under 140 policy indicators.

[5] UNHCR, *Global Appeal 2011 Update*, p. 3. Accessed via http://www.unhcr.org/4cd913ab9.pdf (10 January 2011).

[6] *Ibid.*

[7] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. iii.

Within these operations, indicators have been referred to as a "global yardstick" through which UNHCR can "objectively compare" the wellbeing of diverse populations.[8] A key initiative, the *Practical Guide to the Systematic Use of Standards and Indicators in UNHCR Operations* (the '*Practical Guide*'), seeks amongst other efforts to standardize data collection within 111 refugee camps under UNHCR control.[9]

Despite the breadth of their coverage, the use of indicators as a tool of governance within UNHCR has not yet generated sustained critical attention from outside the organization.[10] Given the centrality of indicators to the distribution of resources and ongoing project development within UNHCR, the need for analysis of their structure and use is pressing.

This paper approaches the issue in three parts. Part I, "Proliferation of indicators in global governance", outlines recent scholarship on the implications of the emergence of indicators as a form of governance in the global sphere. Concerns raised are divided into problems of conception, problems of creation, and problems of use. Part II, "Use of indicators within UNHCR", addresses the impetus for the development of indicators within UNHCR, and outlines the current mechanisms in use.

In Part III, the paper assesses the aims of indicator use within UNHCR in light of analytic pitfalls identified in recent scholarship. Analysis is centred on the indicators and methodology outlined in the *Practical Guide*. While recognizing that some form of data gathering and comparative analysis is necessary for UNHCR to perform its functions effectively and efficiently, the paper identifies four key concerns arising from the current indicator framework used in UNHCR operations.

First, the indicators adopted contain structural flaws ranging from the use of inadequate proxies to a requirement that binary judgments be made on complex standards or contested terms.

Second, the substantial burden of data collection in field operations may strain the already stretched capacity of staff to meet their primary obligations to refugee populations.

Third, contested policy choices can be obscured by a process that instead focuses attention on the adequacy or otherwise of data collection. This lopsided approach has the potential to stagnate innovative approaches to meeting needs and consolidate indicator-based proxies as permanent fixtures of operations.

Fourth, the paper queries the capacity of indicators to act as 'accountability mechanisms', in particular testing their ability to provide participatory accountability to UNHCR beneficiaries.

---

[8] *Id*, p. 3.
[9] *Ibid.*
[10] It appears that the only scholar to have critically analysed the use of indicators in UNHCR is K B Sandvik, "On the Relation between Indicators, Law and Rights in Humanitarian Governance" (unpublished, on file from the author).

Proliferation of indicators in global governance

The adoption of indicators within UNHCR operations is a manifestation of a broader turn to indicators as powerful non-deontic instruments of governance in the global sphere. Indicators are emerging as robust tools within international organizations, non-governmental organizations, and private bodies.[11] As examples, the World Bank's *Doing Business* indicators measure and compare the ease of investment in 175 countries, and are credited with having 'inspired or informed' 48 legal reforms globally.[12] Information dissemination through the OECD Programme for International Student Assessment (PISA) has catalysed legislative overhauls in the German education system.[13]

Myriad human rights indicators now seek to measure compliance and violations under treaty regimes with quantitative data collected by State Parties.[14] Even private initiatives have been held responsible for changes in the allocation of resources. The US News University Rankings have been criticized for shaping students' choices despite being based on 'entirely unaudited surveys',[15] and for affecting managerial decisions within faculties in an exercise of 'unintended regulation'.[16]

The perceived benefits of indicators turn on their capacity to improve assessment and accountability within organizations, and to provide key data capable of improving project development and implementation. Indicators are deemed particularly useful to policy-makers charged with measuring efficiency and making complex decisions concerning the allocation of scarce resources. Their use potentially presents a means to hold organizations accountable to multiple stakeholders while providing critical information to shape future decision-making and measure changes in performance through time.

However the exercise of regulatory authority through reliance on quantitative measures also raises questions of accountability, transparency and participation. Recent scholarship has sought to map regulatory aporias in the production and application of indicators, and to scrutinize the implications of indicator use for key stakeholders.[17] This inquiry can be understood as a limb of the global administrative law project and is intimately linked to concerns regarding the accountability, transparency and participation in regulatory decision-making on the global level.[18]

---

[11] See K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*, p. 1.

[12] See K E Davis & M B Kruse, "Taking the Measure of the Law: The Case of the *Doing Business* Project" (2007) 32(4) *Law & Social Inquiry* 1095, p.1119.

[13] See A von Bogdandy & M Goldmann, "The Exercise of International Public Authority through National Policy Assessment: The OECD's Pisa Policy as a Paradigm for a New International Standard Instrument" (2008) 5 *International Organizations Law Review* 241.

[14] See M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights" (2008) *IILJ Working Paper 2008/12*, p. 2.

[15] See D Segal, "Is Law School a Losing Game?" *New York Times* (9 January 2011), Business 1, p. 6.

[16] M Sauder & W Espeland, "Unintentional Regulation: How Measures can become Disciplinary" *paper presented at* Indicators as a Technology of Global Governance, 13-14 September 2010, NYU.

[17] For a list of recent scholarly papers see *supra*, ft 3.

[18] For an overview of the Global Administrative Law project see B Kingsbury, N Krisch & R Stewart, "The Emergence of Global Administrative Law" (2005) 68 *Law and Contemporary Problems* 16.

In order to clarify the scope of the indicators project it is necessary to adopt a clear definition of terms. This paper adopts with reservations the frame proposed by Davis, Kingsbury and Merry, which characterizes an indicator as:

> [A] named, rank-ordered representation of past or projected performance by different units that uses numerical data to simplify a more complex social phenomenon, drawing on scientific expertise and methodology. The representation is capable of being used to compare particular units of analysis (such as countries or persons), and to evaluate their performance by reference to one or more standards.[19]

While this definition captures the general structure and usage of indicators, a further distinction could be made between indicators that are generated by external organizations as a means to monitor a particular organization or set of policies, and those indicators that are internally framed or generated by the monitored institution itself. Drawing a distinction between indicators that are self-generated and those that are generated by third parties also focuses attention on the distinct issues that may arise in each case.

The Davis, Kingsbury and Merry model is directed towards indicators that are generated by third parties and made publicly available – which may adopt rank-ordering both for reasons of public accessibility, and to foster compliance over time via the mobilization of shame. In contrast, internally generated indicators are largely in-house tools designed to assist in planning, coordination and monitoring of projects. Data collected under these indicators remains firmly within the control of the monitored organization. Maintaining control over data enables the organization to address exposed operational failures without attracting wide publicity or exacerbating pressure from donors that have vested interests in particular regions or operations. This lack of transparency can however raise its own concerns, including questions of accuracy and a lack of peer review accountability.

The confidentiality offered by internally generated indicators explains certain structural differences between these indicators and those generated by third parties. Although the results of internally generated indicators can be released to provide accountability to donors or to publicize the results of effective programs, the subjects of measurement are unlikely to be ranked. Indicators adopted by UNHCR are not rank-ordered, although they are otherwise susceptible for analysis under the Davis, Kingsbury and Merry model.[20]

When segments of data drawn from indicators are released in the annual UNHCR *Global Appeal* or *Global Report*, the ordering is based on an alphabetical country or regional listings (or aggregate data on binary questions that conceals individual countries' status) rather than rank.[21] Similarly, the Sphere Project's Minimum

[19] K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*, p. 2.

[20] A similar approach to non-rank ordered data has been taken in work on Sphere and HAP indicators and benchmarks in the humanitarian space. See M L Satterthwaite, "Indicators in Crisis: Rights-based Humanitarian Indicators in Post-Earthquake Haiti" (2011) (forthcoming), p. 5, ft 25.

[21] Aggregate data drawn from indicators has been published in the UNHCR *Global Appeal 2010-2011*, and the UNHCR *Global Appeal 2011 Update*. Additional data on the breakdown of populations of

Standards in Disaster Response, which are framed externally but collected and monitored internally by humanitarian organizations, eschew the rank ordering of publicly released data.[22]

This hesitance to rank internally generated data is significant, in that it removes from indicators an intuitive judgment regarding the comparative quality of measured phenomena, such as the level of protection assistance between camp situations or various resettlement State policies. In the refugee context, this aversion to publicly available comparative ranking can also be explained by the delicate relationships in place between UNHCR and States that are variously stakeholders, donors to UNHCR operations, or host governments to UNHCR camp operations.

The remaining aspects of the Davis, Kingsbury, and Merry definition may be adopted without caveat. The definition directs attention to three key aspects of indicators: their purported scientific objectivity, susceptibility to comparison, and capacity for iterative evaluation. Each of these features is analyzed below through a critique of indicators at the key points of their conception, creation, and use.

Problems of conception

Both internally and externally generated indicators can be understood as a discrete form of knowledge, premised on the reorganization of qualitative information open to interpretation into a technical, objective and unassailable numerical order. Indicators involve simplification of complex data and the reconstruction of politically tinged conclusions as objective 'facts'. The transformation of information about the world into quantitative data involves 'uncertainty absorption' whereby nuance and caveats are elided, and premises shrouded, through the perceived clarity of numerical form.[23] As Merry notes, "[t]he political process of judging and evaluating is transformed into a technical issue of measurement and counting."[24]

The discourse of scientific measurement and technocratic expertise carries implicit assumptions of impartiality and accuracy. An indicator's purported objectivity is therefore key to its capacity to alter institutional behaviour. Even those critical of an unquestioning use of indicators as regulatory tools recognize the potential power of statistical rarification. Despite a wary approach to indicator misuse, Satterthwaite and Rosga note that debates about their centrality "may provide advocates with new

---

concern are recorded under a separate entry for each country. *Global Appeal* reports are accessible at http://www.unhcr.org/pages/49c3646c4b8.html.

[22] The Sphere Project aims to define and promote standards for responses by global actors to the plight of those affected by disasters. The standards are a compilation of best practices in disaster response as outlined in the Humanitarian Charter and Minimum Standards for Disaster Response (the 'Sphere Handbook'): see http://www.sphereproject.org/content/view/91/58/lang,english/ (accessed 29 December 2010). For a useful overview of the creation of Sphere and its use of indicators see M L Satterthwaite, "Indicators in Crisis: Rights-based Humanitarian Indicators in Post-Earthquake Haiti" (2011) (forthcoming), pp. 22-28.

[23] K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*, p. 8.

[24] S Merry, "Measuring the World: Indicators, Human Rights and Global Governance", prepared for presentation at the American Society of International Law Panel on Indicators (25-27 March 2009), pp. 9-10. Accessed via http://www.iilj.org/research/IndicatorsProject.asp (2 November 2010).

opportunities to use the language of science and objectivity to hold governments to account."[25]

This 'trust' in the objectivity of data involves transference of legitimacy. The technical expertise required to build an indicator and its capacity to recast complex data into numeric form may blind the consumer to the interpretative premises of the project. An indicator can therefore be characterized under the German notion of *verdoppelung*, whereby "an instrument of knowledge is reconceived as the object of knowledge."[26]

Problems of creation

The creation of an indicator requires first that decisions be made regarding the terms and scope of the indicator. Subsequently, the most appropriate form of data collection is chosen. The measured phenomena will often be framed as a proxy for broader concerns – for example, the percentage of girls attending primary school may be used as an indicator of gender equity or educational standards.

As Davis, Kingsbury and Merry note, choices in framing an indicator involve implicit assumptions regarding the "appropriate standards against which to measure societies and the appropriate ways in which to measure compliance with those standards".[27] Such choices require that an analytic link be drawn between abstract outcomes and measurable phenomena. The framing of the data to be measured is therefore a deeply contested decision that requires a high level of analysis and broad participatory input from relevant stakeholders and beneficiaries.

Once a proxy is chosen for measurements, wranglings involved in its framing are obscured by concerns regarding the accuracy of data collection. The questions of who is collecting data, where, how often, and with what resources become crucial. Where indicators are internally generated and the monitored organization is responsible for collecting data, it may have an interest in obscuring or falsifying negative results.

Further, indicators are a means by which central offices can control exercises of power by the periphery. Regardless of whether the resulting data is released publicly or used internally to determine the effectiveness of programs, there may be incentives for collection officers to downplay or couch certain failures if they are perceived to reflect poorly on sectoral management.

Cross-checking and oversight of both internally and externally generated indicators is to be preferred if feasible. Collection of data from multiple sources, or different sources over time, may compromise standardization and hinder an accurate measure of changing circumstances. Existing information resources can vary greatly from state to state, hindering neutral comparative analysis. Particularly in fragile states, it may

[25] M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights" (2008) *IILJ Working Paper 2008/12*, p. 5.
[26] On *verdoppelung* see M Koskenniemi, *The Gentle Civilizer of Nations: The Rise and Fall of International Law 1870-1960*, Cambridge University Press, 2001, p. 242.
[27] K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*, p. 2.

be difficult to access governmental information or speak to subjects. If samples are involved, questions arise as to whether they are representative or skewed.

Furthermore, it may be difficult to standardize teams of researchers so that they are equipped with equal technical and linguistic skills over a number of locations. As an example, the collection of data on gender-based violence cannot be achieved solely through hospital and clinic records, but a more holistic approach will prove difficult if researchers are not skilled in collection, provided with interpreters and psychological training, and given sufficient time and resources to form relationships with individual subjects.

Debates surrounding the data underlying indicators have also emerged in the corporate governance field, in particular in the development of metrics to assess the governance of public companies.[28] A recent example is the controversy surrounding the Antidirector Rights index, which had been applied as a comparative measure of shareholder protection in numerous studies.[29] Holger Spamann's study, which tested the accuracy of the index values in several countries (based solely on the accuracy of the data rather than querying the original variables chosen), found that 33 of the 46 observations in the index were inaccurate and required revision.[30]

Problems of use

Once an indicator is finalized and data is collated, attention turns to questions of use. A central feature of indicators is their comparative function.[31] In reducing qualitative information to key, bounded questions, indicators enable the comparison of phenomena in fundamentally different settings. Comparisons are inherent in indicators that organize their data in a rankings system, but are also possible in regimes that organize numerical results by country or regional location.

The power of the comparative function is closely tied to issues of objectivity and expertise. An absence of regulatory control at earlier stages of development may result in data that is contestable but perceived as neutral. The allocation of resources from a fixed budget amongst numerous beneficiaries or projects inevitably involves complex political and ethical judgments regarding needs. Indicators help reframe these political judgments as technical operations. Reducing the space of contestation may insulate decisions from criticism through the impenetrability of the technical premises underlying the indicator.

[28] For a clear overview of these developments, see L A Bebchuk and A Hamdani, "The Elusive Quest for Global Governance Standards" (2009) 157 *University of Pennsylvania Law Review* 1263.
[29] See H Spamann, "Law and Finance Revisited", *Harvard Law School John M. Olin Center Discussion Paper No. 12* (2/2008).
[30] *Id*, pp. 3, 14. Metrics of the study have also been challenged on the basis of they fail to disaggregate data to take account of their ownership structures – L A Bebchuk and A Hamandai argue that any governance metric that purports to apply to companies regardless of ownership structure is "bound to miss the mark": L A Bebchuk and A Hamdani, "The Elusive Quest for Global Governance Standards" (2009) 157 *University of Pennsylvania Law Review* 1263, p. 1263.
[31] Sattherwaithe & Rosga have gone so far as to state that indicators "are assumed to be valuable only insofar as they are cross-nationally comparable": M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights" (2008) *IILJ Working Paper 2008/12*, p. 5.

As a crude example in the refugee context, the choice between delivering additional aid shipments to Sudan or Congo can be justified on the basis of an 'objective' and 'scientific' study in a way that distances the decision made from institutional preferences and the proclivities of the decision-maker.[32] In doing so the technical vocabulary of the indicators assists to defuse criticism of operational choices by the public and by stakeholders.

'Piggy-backing', whereby an indicator developed in one field is used as part of the data collected to support the creation of a new indicator, is also becoming increasingly common. As an example, a UNHCR Urban Programme Indicator, "percentage of asylum seekers/refugees below the national poverty line" directs UNHCR Field Officers to use World Bank indicators as a data source for reporting.[33]

As a practical matter piggy-backing is often the most pragmatic course of action, in that it makes use of existing data rather than expending resources and energy on repeating collection efforts. However it also leaves an indicator open to importing methodological biases or collection errors associated with the initial indicator. To reduce the risk bias or misrepresentation in the new indicator, such practices must only be undertaken after independent analysis of both the underlying assumptions of the primary indicator and the scientific validity of its collection methods.

Such assessment is particularly important if creation of the new indicator will have the effect of crystallizing a 'soft' regulatory indicator into an aspect of a hard mode of governance. Failure to assess the value of the primary indicator could result in compounding underlying problems at the point of framing, data collection or compilation.

Finally, indicator use cannot be understood simply as a reaction to a single iteration of data. Indicators are conceived not merely as tools of assessment but as tools of improvement. The re-assessment of indicators over time both measures changes in information and encourages internal reform that will be reflected in improved quantitative results. Sauder and Espeland refer to this occurrence as 'reaxivity' – the soft compliance pull of data on the behaviour of measured entities over time.[34] Conceiving of indicators on a temporal axis also uncovers other issues. These include the ossification of particular modes of operation that are aimed at improving performance under the proxy of the indicator rather than meeting the demands of the

[32] Satterthwaite & Rosga argue that standards have the potential to "depoliticize choices that would otherwise be openly contested in the public sphere": See M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights", (2008) *IILJ Working Paper 2008/12*, p. 48 (citing S Merry, "Measuring the World: Indicators, Human Rights and Global Governance", prepared for presentation at the American Society of International Law Panel on Indicators (25-27 March 2009), pp. 9-10. Accessed via http://www.iilj.org/research/IndicatorsProject.asp (2 November 2010); B Jacobsson, "Standardization and Expert Knowledge" in *A World of Standards* (N Brunsson et al eds.) Oxford University Press, 2000, pp. 40-49.

[33] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Urban Programme Indicator 25.

[34] M Sauder & W Espeland, "Unintentional Regulation: How Measures can become Disciplinary" paper presented at *Indicators as a Technology of Global Governance*, New York University (13-14 September 2010).

underlying mandate,[35] and the development of institutional strategies to create the veneer of progress.

These tendencies may develop both in externally generated indicators, driven by the effect of public pressure and the mobilization of shame, and in internally generated indicators, in an attempt to demonstrate strong performance at a departmental or field level to superiors. Taking a long-term view of indicators also raises the possibility of unnoticed mutation within named and established indicators, whereby the "name's constant" can "mask changes over time in the indicator itself".[36]

The concerns outlined above reflect a cautious approach to indicators that has become increasingly widespread as indicators have gained popularity as a tool of global governance.[37] Such an approach sits in contrast to characterizations of indicators in policy documents, which cast indicators as tools for *enhancing* accountability, transparency and participation. In the humanitarian sector in particular, indicators are widely considered to be capable of promoting the efficiency and effectiveness of projects, satisfying stakeholders including donor states, and enhancing accountability to beneficiaries via participatory mechanisms.[38] These contrasting accounts are investigated below in light of the specific mechanisms in place in UNHCR's current operations.

Use of indicators within UNHCR

At the outset, it should be recognized that there is a serious need for accurate, targeted, and regularly updated data within UNHCR operations. Use of standards and indicators in UNHCR has its roots in an effort to enhance the efficiency, consistency and oversight of operations. The mid-1990s saw a move towards a managerialist and outcomes-driven approach in humanitarian assistance projects, typified by the implementation of Results-Based Management ("RBM").[39] RBM is driven by the principle "that impact can be measured and results objectively evaluated".[40]

---

[35] On this point see M L Satterthwaite & A Rosga, "The Trust in Indicators: Measuring Human Rights" (2008) *IILJ Working Paper 2008/12*, p. 53.

[36] K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series*, p. 4.

[37] This caution is now extending beyond academia to encompass professionals working in the development field. A recent initiative at the Institute of Development Studies, "The Big Push Back", brought together 70 development officials to develop strategies to counter an "audit culture" in which "[f]unding agencies are increasingly imposing extraordinary demands in terms of reporting against indicators of achievement that bear little relation to the manner and possibilities that development activities have for supporting social transformation": see Institute for Development Studies, "Development Professionals Launch "Big Push Back to counter audit culture" (1 October 2010), Accessed at http://www.ids.ac.uk/go/news/.

[38] See M L Satterthwaite, "Indicators in Crisis: Rights-based Humanitarian Indicators in Post-Earthquake Haiti" (2011) (forthcoming), p. 11 (relying on surveys with staff working in the humanitarian sector in Haiti).

[39] Although the use of indicators in UNHCR operations is novel, the organization has relied on certain forms of statistical data since its inception. Article 35(2) of the 1951 *Convention Relating to the Status of Refugees* ("the 1951 *Convention*") obliges Contracting States to provide statistical data to the Office of the High Commissioner relating to the condition of refugees, the implementation of the Convention, and any laws and decrees relating to refugees upon request. An analogous provision is included in the 1967 *Protocol Relating to the Status of Refugees* at Art II(2)(a)-(c). While the 1951 *Convention* and

This impulse towards reform was driven by serious and endemic accountability flaws highlighted in this period, as the organization's practices were called into question in a series of audits and subsequent reports by the UN internal watchdog, the Office of Internal Oversight Services (the "OIOS"). The 1995 Report of the Secretary-General on the activities of the OIOS noted that audits of UNHCR operations in 28 member States revealed:

> non-compliance with agreements with UNHCR, lack of accountability for expenditures, absence of optimal procurement arrangements, substandard property control and ineffective distribution of relief items to the refugee caseloads.[41]

Causes of inefficiencies were linked to inadequate staffing and facilities in partner organizations, volatile circumstances in field offices, and "poor management and control on the part of UNHCR".[42] The OIOS recommended that UNHCR enhance its control over partners, monitoring arrangements, transparency and remedial strategies.[43] A subsequent report in 1996 held that UNHCR had improved its effectiveness in some areas but continued to show weaknesses in its finances, dealings with implementing partners and the incapacity of central agencies to monitor activities in the field.[44]

Audits found that funds awarded to implementing partners were not maintained in separate bank accounts or accounting ledgers, hampering an accurate assessment of actual project expenditure, and that implementing partners "could not always account for substantial project expenditure reported to UNHCR".[45] Lack of external oversight and the high stakes of refugee operations created opportunities for corruption by unscrupulous administrators and third parties. An investigation carried out at the request of UNHCR by OIOS into its Nairobi operations exposed a multi-tiered scheme in the late 1990s to extort money from refugees in exchange for providing them with permanent resettlement in third countries.[46] Over 70 people were implicated in the scheme, including UNHCR employees and administrators of affiliated NGOs.

1967 *Protocol* are concerned with the collection of statistics by State Parties themselves, the recent use of indicators relies on internal collection of data.

[40] UNHCR, "Update on Global Needs Assessment (GNA) and Results-Based Management (RBM)", *Informal Consultative Meetings* (6 February 2009).
Accessed at http://www.unhcr.org/499180492.html.

[41] UNGA, *Report of the Secretary-General on the activities of the Office of Internal Oversight Services*, A/50/459 (2 October 1995), [44]-[45]. UNHCR projects were audited in Armenia, Bangladesh, Belize, Benin, Botswana, Burundi, Chad, China, the Congo, El Salvador, Georgia, Ghana, Guinea, Hungary, India, Iraq, Kenya, Laos, the Libyan Arab Jamahiriya, Namibia, Somalia, Thailand, Togo, Uganda, Venezuela, the former Yugoslavia, Zaire and Zambia.

[42] *Ibid.*

[43] *Id*, para. [46].

[44] OIOS, *Triennial review of the implementation of recommendations made by the Committee on Programme and Coordination (CPC) at its 32nd session on the evaluation of the Office of the United Nations High Commissioner for Refugees (UNHCR)*, E/AC.51/1996/4 (21 March 1996), paras. [22]; [26]-[27]; [34]-[35]. See also UNGA, *Report of the Secretary-General on the activities of the Office of Internal Oversight Services*, A/51/432 (30 September 1996), paras. [44]-[49].

[45] UNGA, *Report of the Secretary-General on the activities of the Office of Internal Oversight Services*, A/51/432 (30 September 1996), paras. [44]-[49].

[46] Irin News, "Kenya: UNHCR head accepts Nairobi corruption report" (28 January 2002). Accessed via http://www.irinnews.org/report.aspx?reportid=29962.

These specific instances of mismanagement and corruption stand alongside endemic problems of measurement and aid delivery in refugee situations. Accurate estimates of the amount of aid needed for refugee communities, or even the size of populations within camps, are undermined by unpredictable population flows and incentives for refugees themselves to inflate official population figures. Higher population estimates result in increased aid deliveries, which can then be used to supplement a family's diet or to barter on the open market. Kibreab identifies a range of strategies used by refugees to distort camp statistics and increase aid provision:

> collusion with host government officials; inflating numerical sizes of families for purposes of registration; withholding of information on deaths; registration in different sites in order to obtain multiple ration cards; double or even triple registration; splitting of families between different camps (the latter may even happen during repatriation); and the exhibition of physical helplessness in the presence of aid givers.[47]

An extreme example of population inflation occurred in the Tog Wojaale refugee camp in north-west Somalia – the official estimate of 87,000 people was found to exceed the actual population by 55,000 when an official census was finally undertaken in 1987.[48] In the aftermath of the 2010 earthquake in Port-au-Prince, Haiti, reports surfaced of "ghost camps" set up by impoverished Haitians who pose as Internally Displaced Persons (IDPs) during the day to obtain aid and return to their homes in the evening.[49] There have also been cases of host States themselves inflating population statistics in order to attract additional aid.[50]

While the rise of RBM within UNHCR was catalyzed by the specific concerns outlined above, its implementation was also influenced by broader developments in the humanitarian field.[51] As Satterthwaite has outlined in her work on indicators in the humanitarian sector, recognition of the need for stronger accountability mechanisms and efficiency controls stemmed from a series of failures in the late 1980s and early 1990s, epitomized by the Srebrenica massacre in "safe zones" ostensibly protected by the UN Protection force (UNPROFOR), and humanitarian actors' incapacity to stem violence during the Rwandan genocide in 1994.[52]

A quest for increased coordination between agencies, and greater accountability both to donors and beneficiaries, led to the development of the Active Learning Network

---

[47] G Kibreab, "Pulling the Wool over the Eyes of the Strangers: Refugee Deceit and Trickery in Institutionalized settings", (2004) 17 *Journal of Refugee Studies* 1, p. 1.

[48] *Id*, p. 3, citing studies by D Jamieson, "The Demise of Tug Wajele" (1987) 40 *Refugees* 30, p. 30 that the false figures (which in turn led to a gross surplus in the provision of aid granted to the camp) were generated by refugees crossing the nearby border back into Ethiopia and then returning as "new arrivals".

[49] See M L Satterthwaite, "Indicators in Crisis: Rights-based Humanitarian Indicators in Post-Earthquake Haiti" (2011) (forthcoming), pp. 70-73.

[50] See M F N Franke, "Refugee Registration as Foreclosure of the Freedom to Move: The Virtualisation of Refugees' Rights within Maps of International Protection" (2009) 27 *Environment and Planning D: Society and Space* 352, p. 355, citing C Dolan & J Large, "Evaluation of UNHCR's repatriation and reintegration programme in East Timor, 1999-2003", EPAU/2004/02 (2004).

[51] See discussions of the Sphere Project and other humanitarian RBM measures in UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), pp. vii; 14-15.

[52] See M L Satterthwaite, "Indicators in Crisis: Rights-based Humanitarian Indicators in Post-Earthquake Haiti" (2011) (forthcoming), pp. 17-18.

for Accountability and Performance (ALNAP),[53] and the Sphere Project, a policy initiative launched in 1996 by the humanitarian sector that aimed to increase accountability and efficiency through the publication of technical standards for disaster response.[54] The Sphere Project is particularly significant for its efforts to enhance consistency and quality in the standards adhered to by NGOs while retaining the independence from government oversight necessary to maintain humanitarian agents' neutral status.[55] Satterthwaite views this model of accountability – mapped by adherence to standards rather than obligations to stakeholders – as one of the "hallmarks" of the Sphere project.[56]

Both the endemic problems faced in the "emergency zone" of refugee assistance and the particular crises highlighted in the mid-1990s should be borne in mind when mapping the development of RBM mechanisms and indicators within UNHCR operations. The implementation of standards and indicators in UNHCR operations is one aspect of a project to enhance accountability both for external stakeholders and for beneficiaries, while improving efficiency and streamlining coordination with implementing partners.

Standards and Indicators Initiative

In 2002 UNHCR launched the "Standards and Indicators Initiative" as a means to develop a series of mechanisms that could structure assessment, planning and implementation of operations.[57] Originally restricted to the assessment of conditions within camps, the UNHCR standards and indicators framework has since expanded to the collection of data on conditions in camp settings, urban settings, and returnee areas.[58] The current edition of the *Practical Guide to the Systematic Use of Standards and Indicators* (the '*Practical Guide'*), released in 2006, aims to collect information on 111 camp situations, 107 countries with urban refugees, and 13 reintegration country operations.[59]

The *Practical Guide* - which runs to over 300 pages – contains a detailed outline of 154 separate indicators. It includes a 'how-to' guide for UNHCR officers collecting data, comments on where data should be sourced for each indicator, and notes on potential pitfalls in measurement. Additional specific indicators have been adopted to

[53] *Id*, p. 19.
[54] *Ibid*. The use of indicators within UNHCR is tied to perceived successes of the Sphere project: see UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. 14.
[55] *Ibid.* Satterthwaite notes that self-regulation "was perceived by a sector that defined itself by its independence as a threat from both donor states and host states".
[56] *Id*, p. 20.
[57] UNHCR, "Update on Global Needs Assessment (GNA) and Results-Based Management (RBM)" *Informal Consultative Meetings* (6 February 2009). Accessed at: http://www.unhcr.org/499180492.html.
[58] An initial set of indicators was issued in 2003 together with a *Camp Indicators Report* to record data collected in the field. These indicators were built into the first edition of the *Practical Guide to the Systematic Use of Standards and Indicators in UNHCR Operations,* published in 2004. An additional list of indicators relating to non-camp situations was first circulated to Field Offices in 2004, and are outlined substantively in the 2006 2nd Edition of the *Practical Guide*. See UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. 3.
[59] *Ibid*.

measure outcomes in health and refugee education levels.[60] In its current form the project is still considered to be a "work in progress" towards the creation of a robust "monitoring tool covering UNHCR operations worldwide".[61] Raw data collected under indicators is used for internal programming but not released to the public.[62] However, restricted summations of the data are compiled in a series of publications aimed towards donors and stakeholders including the UNHCR *Global Appeal,* and *Global Report*.[63] The *Practical Guide* notes that quantitative indicators are intended to be complemented with "additional information from qualitative indicators, project-level indicators, detailed sectoral data, and various narrative reports".[64]

UNHCR presents the *Practical Guide* as providing a "sound empirical base for assessing the quality of UNHCR's protection and assistance activities".[65] Three key aims are identified for the use of these indicators. First, indicators are characterized as 'measures of progress' that assess the impact of existing programs through iterative measurement over time.[66] Second, indicators are considered to be a planning tool in that they provide information of relevance to future operations and promote consistency in the distribution of resources.

Finally, the use of indicators is intended to improve accountability within UNHCR operations by promoting a culture of transparency and participation. Stakeholders are diverse; as accountability mechanisms, indicators are directed variously at refugee populations, host States, donors, members of UNHCR's governing Executive Committee and "other parties within the humanitarian community that support and monitor the activities of UNHCR."[67]

Overall responsibility for reporting rests with UNHCR country offices for data collected on a country level, in urban programmes and in returnee areas. UNHCR field officers are responsible for data collected on refugee camps. Data is generated from two sources: existing data collected by external institutions including implementing partners, host government authorities, UN agencies and NGOs, and data collected by UNHCR staff themselves.[68] Staff are encouraged to only collect new data "if it is not already available from another source".[69] Given concerns about the

[60] See respectively UNHCR, "Health Information System: Standards and Indicators Guide" (rev. January 2010), accessed via http://www.unhcr.org/4614ab8e2.html (5 November 2010); UNHCR, "Refugee Indicators 2003: education indicators and gap analysis covering 118 camps in 23 asylum countries" (rev 1. 20 August 2004), accessed via http://www.unhcr.org/cgi-bin/texis/vtx/search?page=search&docid=40e426cd4&query=indicators%20refugee%20education (8 November 2010).
[61] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), pp. vii-viii.
[62] I am grateful to Heinrik Pilgaard, DPSM UNHCR HQ Geneva for clarifying this point.
[63] *Ibid.*
[64] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. vii.
[65] *Report of the High Commissioner for Refugees*, (GAOR, Suppl. No 12 A/60/12), p. 9.
[66] See UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), noting that indicators 'measure signs of change towards the achievement of results' (p. 25); and are "a variable scale on which it is possible to objectively measure different points and that corresponds to, or correlates closely with, variations in the conditions of the refugees and persons of concern." (p. 28)
[67] *Id,* p. vii.
[68] *Id*, p. 46.
[69] *Ibid*.

availability of data and accuracy, more than one source of data may be used in the creation of indicators.[70]

The collation of data from external sources is itself a substantial endeavour. As an example, staff charged with monthly measurement of the crude mortality rate in a camp are instructed to verify data through multiple sources including health partner records, records of community service partners, and refugee committees, as well as engaging with "community health workers or grave watchers specifically to collect mortality data to augment hospital and health centre data", and examining records such as "the number of burial shrouds provided".[71] When planning and conducting surveys, field staff are encouraged to rely on "specialized NGOs, academic institutions, and statistical offices".[72]

Results Framework

A second use of indicators is in UNHCR's *Results Framework*, an internal planning tool that aims to categorize all activities conducted by UNHCR, and links each category to "objectives", "outputs", and "impact and performance indicators".[73] Performance indicators aim to measure progress in achieving of given outputs over time,[74] while impact indicators aim to measure changes in the actual situation of persons of concern over time.[75]

The name 'impact indicator' contains an implicit assumption that the changes in situation can be directly linked to the projects implemented by UNHCR. This assumption has been criticized by some UNHCR field staff as ignoring the complex interplay of factors that affect the situation of refugees and asylum-seekers, and falsely equating the "protection and well-being of persons of concern with the quality of UNHCR's work".[76]

The indicators used in the *Results Framework* are distinct from those outlined in the *Practical Guide*, although there is some overlap.[77] The *Results Framework* is not released publicly, however individual copies may be requested through the UNHCR Policy and Development and Evaluation Service.[78] While the methodology and

[70] *Id,* p. 47.
[71] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Refugee Camp/Settlement Indicator 4.
[72] *Id,* p. 49.
[73] R Allen & A Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 19. Accessed at http://www.unhcr.org/4cf3ad8f9.html.
[74] *Ibid.*
[75] *Ibid.*
[76] *Id,* p. 21. Allen and Rosi cite the results of interviews with Field Officers regarding the implementation of RBM strategies. The report is based on over 100 interviews with UNHCR staff and implementing partners both at Geneva Headquarters and in seven different Field Offices: *Id,* p. 6.
[77] *Id,* p. 21, citing comments from Field Officers noting a lack of continuity between the two documents. However, results of indicators published in UNHCR *Global Appeal 2010-2011* and the *Global Appeal 2011 Update* demonstrate some overlap with indicators measured in the Practical Guide, particularly those addressing education, Sexual and Gender-Based Violence Strategies, and domestic legislative frameworks.
[78] I am grateful to Jeff Crisp and Angela Li Rosi at the Policy and Development Evaluation Service ('PDES') for offering a copy of the *Results Framework* upon request.

structure of the indicators is less accessible than those in the *Practical Guide*, aggregate data from the *Results Framework* indicators have been published in recent editions of the UNHCR *Global Appeals*, enabling tentative findings to be made on the indicators' intended scope and use.

Global Strategic Priorities

A third use of indicators is in the context of UNHCR's Global Strategic Priorities, a set of objectives that are set annually and published in UNHCR's annual *Global Appeal*.[79] As noted above, indicators and priorities listed in the *Global Appeal* are drawn from the *Results Framework*.[80] The Global Strategic Priorities replace the Global Strategic Objectives, which were in place from 2004-2007. According to UNHCR, the Priorities:

> build upon the global strategic objectives that UNHCR used in previous years to strengthen Results-Based Management. They will help the Office to address the needs of refugees and others of concern by defining a common set of priorities as a blueprint for global operations.[81]

Each priority is linked to an indicator which is purported to measure progress towards achievement of the priority, and a 'description of the current situation', which is based on data from a variety of sources including collations of previous standards and indicator reports.[82] A target to work towards is also listed.

According to the *Global Appeal 2011 Update*, the Priorities are used in part to "guide allocation of resources" and "identify local priorities".[83] As such, the information collected under the Global Strategic Priorities, and conclusions on the current situation in a given territory, are significant to conclusions on funding and resource distribution made in high-level planning.

The Global Strategic Priorities are also significant in that they are one of the few areas where information collected under UNHCR indicators is released to the public. In the *Global Appeal 2010-2011*, UNHCR included alongside priorities a summary of the global "current situation" regarding each aim. As an example, Priority 1.1, "*The rights of persons of concern are recognized in law and in practice. There is adequate administrative capacity to support the implementation of international protection standards"* contains a note that in at least 108 countries, the national legal framework does not adequately meet international protection standards.[84]

No disaggregation is provided to determine which countries are included in the 108 countries with inadequate national legislation. The claim involves a high level of critical analysis and subjectivity. Without access to the reports made in various

[79] On the measurement aspect of Global Strategic Priorities see *Report of the High Commissioner for Refugees*, (GAOR, Suppl. No 12 A/60/12), pp. 9-10.
[80] UNHCR, *Global Appeal 2010-2011*, p. 10. Accessed at http://www.unhcr.org/ga10/index.html.
[81] *Ibid.*
[82] UNHCR, *Global Appeal 2011 Update*, p. 12. Accessed at http://www.unhcr.org/ga11/index.html.
[83] *Ibid.*
[84] *Id*, p. 13.

countries, it is difficult to ensure that the same standards of "adequacy" were used throughout. The conclusion alongside Priority 2.1 that "in at least 78 countries reception conditions are not adequate" raises similar issues.[85] It is not clear how the word "adequate" is defined, whether a consistent definition is adopted throughout reporting, and whether issues of national capacity were taken into account as mitigating factors in the characterisation of the 'reception conditions' in place.

Other uses

Indicators are also used in UNHCR as a measure of the impact of the Millennium Development Goals upon the lives of persons of concern.[86] The 2006 Standards and Indicators Guide lays out causal links between specific goals and UNHCR indicators. As an example, a linkage is made between the MDG target that by 2015 children everywhere will be able to complete a full course of primary schooling, and UNHCR indicators on the "percentage of refugee students enrolled in primary school" and "percentage of refugee school aged population living within reasonable distance from primary school".[87]

Several of these linkages are decidedly indirect – indicators linked to the elimination of gender disparity in primary and secondary education include "percentage of needs met for sanitary materials", "percentage of female members in management committees", and "percentage of female members in food distribution committees".[88] Given the difficulty of demonstrating causal connection between progress on UNHCR indicators and their linked MDG Goals and Targets, these measures may not provide a solid foundation for measuring the fulfilment of broader development goals.

Implementing partners are also called on to use indicators in their work. According to interviews carried out by UNHCR's Policy Development and Evaluation Service, a number of Agreements between UNHCR and implementing partners include additional indicators that are not listed in either the *Results Framework* or the *Practical Guide*.[89]

Indications of progress? An assessment of UNHCR's current indicator model

In light of the complexity of UNHCR operations and past failures in service delivery, it is clear that some means of assessment and measurement is necessary. However the indicator model currently in place has a number of weaknesses. While recognizing that the project is still in a state of development, it is important to assess existing shortcomings given the highly ambitious nature of the program, the lack of cross-checking or external regulation in the collection of data, and the significant implications of indicator reporting for UNHCR project planning, aid distribution and budgeting.

Many concerns outlined below can be addressed and improved upon in future versions of the *Practical Guide*. Others are either difficult to rectify at this stage of the project's development, or inherent in the use of indicators as an operational tool.

[85] See the current situation measure linked to Global Strategic Priority 2.1: "*Persons of Concern are registered or profiled in a manner that enhances protection. Reception arrangements improve their security and access to essential services"*: UNHCR, *Global Appeal 2011 Update,* p. 14.

Even in the latter case, evaluation is useful in generating a more critical attitude towards indicators at the managerial level.

Attention to the shortcomings of indicators will encourage their use alongside other tools of measurement and may help to avoid problems of ossification in the future. For pragmatic reasons of space, analysis is centred on indicators in the *Practical Guide*. Indicators drawn from the *Results Framework* are discussed indirectly where aggregate results have been published by UNHCR in recent *Global Appeal* reports.

Structural flaws

A preliminary analysis of the 154 indicators outlined in the *Practical Guide* raises six basic structural concerns. It is not suggested that each of these concerns is present in every indicator, although some indicators may attract more than one concern. These six concerns are that:

- phenomena measured under the indicators are often inadequate proxies for the underlying issues that are purportedly assessed;
- several indicators call on the data collector to form binary judgements on complex questions, which could lead to inconsistencies in measurement between different country operations;
- a lack of disaggregation in measured phenomena leaves indicators open to misinterpretation or abuse;
- misleading or ambiguous titles to several indicators deflect attention from the actual phenomena measured;
- it is often unclear whether an indicator measures legislative frameworks or actual practice; and
- several indicators call on data collectors to make judgements on vague, subjective standards.

Each of these issues is discussed in turn below.

---

[86] See *Report of the High Commissioner for Refugees*, (GAOR, Suppl. No 12 A/60/12), p. 9.
[87] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. 12.
[88] *Ibid*. While studies are ongoing to determine whether the provision of adequate sanitary materials affects the school attendance of pubescent and post-pubescent women, it is arguable that a clear pattern of causation is yet to be established. For discussion of this issue in the context of Nepal, see E Oster and R Thornton, "Menstruation, Sanitary Products, and School Attendance: Evidence from a Randomized Evaluation", (2011) 3(1) *American Economic Journal: Applied Economics* 91.
[89] R Allen & A L Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 22. Accessed at http://www.unhcr.org/4cf3ad8f9.html.

Inadequate proxies

Perhaps as a result of the quest for specific, comparable phenomena, indicators often bear a tenuous relation to the issues that they are used to measure. As an example, Urban Program Indicator 23 measures the "percentage of female members in asylum seeker/refugee bodies", as an indication of the empowerment of refugee women. However accompanying notes on the indicator state that "women's participation means not only physical presence at meetings but also voicing their concerns and opinions".[90]

The nature of women's actual involvement in meetings, a factor not conducive to quantitative measurement, is obscured by the data collection process. The note therefore appears a blunt acknowledgement of the indicator's inability to truly assess the situation. Similarly, Refugee Camp and Urban Program indicators measure the percentage of refugee students enrolled in classes,[91] a measure that may or may not have a relationship to the number of refugee students that actually *attend* class on a regular basis.

UNHCR Field Officers have further questioned the validity of this measure on the basis that it is linked to the Global Strategic Priority that "boys and girls of concern have equal access to quality primary and secondary education", a question that is largely divorced from statistics on enrollment.[92]

Binary judgments on complex phenomena

A number of the *Practical Guide* indicators call for data collectors to make binary judgements on complex questions with multiple criteria. Country Level Indicator 17 calls for Field Officers to gives a 'Yes/No' response to the question "are rejected asylum seekers being deported/repatriated". The *Practical Guide* directs Field Officers to form this judgement on the basis of UNHCR visits and presence in detention facilities, government sources, human rights agencies, implementing partners, the media, and refugees.[93]

No guidance is given on the number of cases that would warrant a "Yes" answer (it is unclear if one case or ever a handful be sufficient), how a response should measure a situation where there are cases of repatriation but not deportation, or what hierarchy should be given to the data sources in reaching a subjective assessment.

Similar concerns arise in Country Level Indicator 6, "are school diplomas/certificates obtained in the country of origin recognized in host country" (Yes/No), which requires data collectors to look both to legislation and to actual practice, and gives no

[90] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Urban Programme Indicator 23.
[91] See e.g. Urban Programme Indicator 27, "Percentage of refugee students enrolled in Grades 1-6"**;** Urban Programme Indicator 28, "Percentage of refugee students enrolled in Grades 7-12"; and Urban Programme Indicator 29, "Percentage of students with specific needs attending Grades 1-6".
[92] This argument is noted in R Allen & A L Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 18. Accessed at http://www.unhcr.org/4cf3ad8f9.html.
[93] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Country Level Indicator 17.

scope to the definition of 'school diplomas/certificates'.[94] The crude measure of analysis renders these indicators open to the subjective judgment of Field Officers, who may reach very different conclusions on information from similar sources.

Further, the Yes/No standard is liable to render cases deemed 'unusual' or unproven invisible. While the comments section on data report forms enables some discussion of these issues, as qualitative tools they cannot be included in quantitative compilations of the data.

Aggregation of discrete phenomena

A decision to lump together qualitatively different manifestations of phenomena raises questions of whether the resulting data will be useful or comparable across regions. Country Level Indicator 19 purportedly measures the "average processing time between submission of [a Refugee Status Determination] application and final decision (in days)."[95] No disaggregation is made between decisions carried out by UNHCR and national governments, and the notes section of the indicator refers only to appropriate lengths of time for decisions made by UNHCR.

If national decisions are intended to also be included, they are likely to greatly distort the figures given the access of asylum-seekers to judicial review in national appellate Courts or federal courts exercising their original jurisdiction. It is not clear what value the measurement holds if it does not distinguish between these two very different systems of decision-making.

Similarly, Urban Program Indicator 26 conflates asylum seekers and refugees in the measurement of the percentage of individuals who are employed or self-employed, despite the fact that numerous countries adopt different legislative restrictions on working conditions for asylum seekers and refugees.[96]

Misleading names

As Davis, Kingsbury and Merry have noted, the name chosen for an indicator is crucial to how data will be interpreted.[97] Over time, a standard name for an indicator may mask changes to the undergirding methodology, rendering comparative analysis problematic. However a misleading name can also be problematic at the first iteration of measurement. Close reading of the methodology outlined in the *Practical Guide* often reveals a distinctly different phenomena is measured than that suggested by the indicator's name.

In the measurement of Urban Program Indicator 33, "percentage of refugees who repatriated", Field Officers are instructed not to include "spontaneous returns", but

[94] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Country Level Indicator 6.
[95] *Id,* Country Level Indicator 19.
[96] As a comparison, Country Level Indicator 8, "Do all asylum-seekers/refugees have a right to engage in gainful employment" requires data collectors to give separate answers for asylum-seekers and refugees: *Id,* Country Level Indicator 8.
[97] K E Davis, B Kingsbury & S E Merry, "Indicators as a Technology of Global Governance" (2010) *IILJ Working Paper 2010/2: Global Administrative Law Series,* p. 4.

only those who had previously filled out a Voluntary Repatriation Form before returning to their home country.[98] Such a measurement clearly elides a large number of repatriations which, while less easy to measure, should nonetheless be acknowledged in the framing of UNHCR operations.

Even where the title of the indicator appears clear, it may be warped in the eventual publication of data. Country Indicator 1, "Have border authorities been clearly instructed to refer asylum-seekers to the competent authority", is measured solely on a country's legislation and does not measure actual practice.[99] However the *Global Appeal 2010-2011* recasts this indicator as a measure of practice rather than legislation, stating "[i]n at least 63 countries, border authorities do not systematically refer asylum seekers to competent authorities."[100] The indicator itself gives no guidance on border authorities' actual behaviour, rendering the statement in the *Global Appeal* inaccurate.

Notes to data collectors for some indicators seem to contradict the unit of measurement entirely. Country Level Indicator 4 asks "do all asylum seekers/refugees have a right to engage in gainful employment". However the notes state:

> In countries where refugees are, for example, only granted the legal right to engage in gainful employment some time after their stay has been regularized (made lawful), please reply "Yes" to this indicator and provide a brief explanation.

The advice effectively sanctions a record that asylum seekers (i.e., those that are not yet 'regularized') have a right to work that in fact they do not possess. Such advice has the potential to distort final findings and to generate inaccurate results. While it is to UNHCR's credit that a comments box is included to clarify the chosen measure, it is not clear how (if at all) these comments are incorporated into final, comparative analyses of data across countries and operations at a managerial level.

Data sources

On the face of the indicators in the *Practical Guide*, it often unclear whether actual practice or a country's legislative framework forms the basis of analysis. Notes for the indicators "do asylum seekers have access to [compulsory and] free primary education",[101] "Do refugees and asylum seekers have access to emergency and primary health care services without discrimination",[102] and "are births of non-refugee stateless children properly registered"[103] restrict the data source to government sources, although intuitively the words "access" and "properly" would suggest the measurement of actual practice.

---

[98] A similar issue emerges in Country Level Indicator 6 "Number of Refugees Killed", which gives instructions to data collectors not to include death penalty cases or death from natural causes as a result of the displacement (such as starvation): see UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Country Level Indicator 6.
[99] *Id*, Country Level Indicator 1.
[100] UNHCR, *Global Appeal 2010-2011*, p 13. Accessed at http://www.unhcr.org/ga10/index.html.
[101] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Country Level Indicator 5.
[102] *Id*, Returnee Area Indicator 18.
[103] *Id*, Country Level Indicator 12.

Conversely the indicator "are applicants permitted to remain in the country while an appeal is pending" calls on Field Officers to refer to legislation *and* actual practice, despite the legal connotation of the word "permitted".[104] If Field Officers are careful to follow the data source instructions of each indicator closely, and those charged with managerial decisions on the distribution of funds and project planning are attuned to these differences and well versed in the methodology of each of the 154 indicators, this problem would not raise serious concerns. Such an assumption should not be made lightly given the information overload faced by managerial decision-makers and the significant burden of reporting placed on UNHCR staff.

Indicators that encourage data collectors to look to multiple data sources, such as the Refugee Camp/Settlement Indicator on crude mortality rates discussed above,[105] raise difficult comparative questions. Choices made by data collectors can result in a horizontal comparative exercise between camps where collected data is drawn from fundamentally different instruments, each with its own internal shortcomings and biases. Variety manifests itself as a lack of consistency in measurement that can undermine the technical capacity of the indicator and ultimately the 'objective' function that it is intended to fulfil in the planning process.

Vague standards

Vague standards also pose concerns in the collection of consistent, comparable data. Field Officers are required to make subjective assessments on the meaning of terms such as "credible",[106] "adequate",[107] and "received support".[108] UNHCR does attempt in many cases to render terms open to global assessment by taking into account regional needs, however the subjective nature of these assessments may result in data that is devoid of any useful comparative function. A key example is the definition given for "older people", as typically "60 or over", with the qualification that "the application of this policy will respect factors such as life expectancy and cultural norms that differ from region to region."[109]

Initial evaluations by Field Officers into their role of reporting on indicators show that staff perceive this issue to be particularly problematic. In the PDES 'snapshot' report on the implementation of RBM, published comments made by interviewees include "what do 'enhanced self-reliance,' or 'adequate dwelling' actually mean for the

[104] *Id*, Country Level Indicator 16.
[105] *Id*, Refugee Camp/Settlement Indicator 4. See above, pp. 13-14.
[106] See e.g. Country Level Indicator 2 "are there credible reports of refoulement". The only clarification of the word 'credible' is that "reports must be considered credible by UNHCR": *Id*, Country Level Indicator 2.
[107] See Refugee Camp/Settlement Indicator 53 "Percentage of households with adequate dwellings". In an acknowledgement of the difficulties of measurement, the indicator notes state that "universally adequate shelters are difficult to define": *Id*, Refugee Camp/Settlement Indicator 53.
[108] See Urban Program Indicators 12-13: "Percentage of SGBV cases who [*sic*] received support". No specific definition of "support" is given in the notes, although Officers are asked to include a comment if "the survivor victim is proceeding with legal redress or seeks other support, for example psycho-social and medical support": *Id*, Urban Programme Indicators 12-13.
[109] *Id*, Urban Programme Indicator 6, "Percentage of older people with specific needs assisted".

purposes of measurement?",[110] "is an adequate dwelling in Somalia the same as an adequate dwelling in Serbia?",[111] and "for the indicator '# of persons allowed to achieve enhanced self reliance' what does it mean to be 'allowed' and who makes that judgement?"[112] A lack of consistency in these determinations, or a lack of openness regarding the 'sliding scale' methodology of a term such as 'older people' or 'adequate' undermines the comparative function of these indicators.

Methodological concerns: the burden of data collection

The methodology of data collection is crucial to the integrity of indicators both as units of comparative analysis and as tools of measurement over time. Concerns emerge from the structure of collection developed by UNHCR. The Standards and Indicators project envisages that UNHCR Country Office Staff and Field Officers are responsible for collecting data on 154 separate indicators at least annually (and often more regularly).

While not all indicators are intended to be collected in all operations, over 70 indicators are required to be continuously assessed within refugee camps alone. The capacity of UNHCR staff to collect this data accurately and consistently is questionable.

First, inherent hindrances to the collection of accurate data exist in the refugee context. UNHCR explicitly recognises that "obtaining 100% accurate data" is problematic, instead urging that efforts be made to ensure it is "as accurate as possible under the circumstances".[113] In a blunt example, field notes for mortality rate measurement in the *Practical Guide* state that "because there is a tendency to under-report deaths and over-report population, special care should be taken to obtain results that are as accurate as possible".[114] No additional guidance is given on how the intrepid researcher should achieve this aim.

Second, UNHCR staff are charged with the collection and compilation of data on top of their existing duties. All 24 Country Level indicators are required to be measured on an annual basis. For the Urban Program, 35 of the 36 indicators are also measured annually, with only the percentage of asylum seekers receiving income being measured on a bi-annual basis. Of the 29 indicators of returnee areas, 27 are measured annually, while the crude mortality rate is to be measured monthly and the percentage of returnees earning income measured bi-annually.

The majority of indicators are reserved for refugee camp and settlement situations. Of the 71 listed indicators, 10 are measured monthly, 49 are measured at least annually, 1 is measured quarterly, 6 are measured bi-annually, and 4 are measured on the basis of food distribution cycles, reporting periods, or seasonal requirements.

---

[110] R Allen & A Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 21. Accessed at http://www.unhcr.org/4cf3ad8f9.html.
[111] *Ibid.*
[112] *Ibid.*
[113] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. 45.
[114] *Id*, Refugee Camp/Settlement Indicator 4.

The measurement of each individual indicator involves a discrete analysis that may include collection of multiple data sources as well as critical engagement with their methodology, surveys, participatory meetings with refugees, and scientific testing. Even straightforward measurement tasks are apt to tax the resources of Field Officers. A single indicator for Refugee Camps/Settlements, Indicator 46, measures the "percentage of families with latrines". The explanatory notes for data collectors state:

> it is recommended to physically count household latrines in use twice a year, to add monthly new household latrines and to subtract full or collapsed latrines.[115]

The acute optimism of such an instruction is clear in light of the fact that the current ratio of UNHCR staff to persons of concern is 1: 5040,[116] and that in Kenya, 310 staff members are charged with the protection of 447,500 refugees.[117]

UNHCR evaluations on the use of indicators show that practical problems have arisen in the collection of data. An internal account of data collected under the 2003 Camp Indicator Report notes that the quality of reported data "varies greatly", as a result *inter alia* of "unfamiliarity of field staff with the indicators requested, lack of access and insecurity in refugee camps, the nature of UNHCR education programs and the duration of the refugee situation".[118] These problems have not been resolved in the latest iteration of indicators. Interviews undertaken in 2010 by PDES with Field Staff to monitor implementation of new systems showed:

> For some, [the indicators] were considered to be unworkable, and so they were ignored. Others took a more diligent approach, and tried hard to fill in baseline and target data for the indicators. Even so, many struggled with this task.[119]

The complexity of the task of data collection now placed on Field Officers is arguably inconsistent with their existing obligations. In a seminal report on UNHCR culture supported by UNHCR's Evaluation, Policy and Analysis Unit, Wigley argues that the primary task of UNHCR – namely to safeguard the rights and well-being of refugees and to lead and coordinate international action for the world-wide protection of refugees and the resolution of refugee problems – is often subverted by a focus on secondary or auxiliary tasks required to meet organizational goals.[120] Wigley writes:

> UNHCR has a tendency to behave as though its primary purpose is, for example, to create reports, arrange staff movements and keep itself funded, rather than that these are all activities that occur only as a

[115] *Id*, Refugee Camp/Settlement Indicator 46.
[116] UNHCR, *Global Appeal 2010-2011*, p. 3. Accessed at http://www.unhcr.org/4b03d2399.pdf.
[117] *Ibid*, pp. 75-76.
[118] UNHCR, "2003 Refugee Indicators from the Camp Indicator Report: Education Indicators and Gap Analysis covering 118 Refugee Camps in 23 Asylum Countries Based on Initial Data from the Camp Indicator Report" (20 August 2004), p. 3. Accessed via http://www.unhcr.org/40e426cd4.html (10 November 2010).
[119] R Allen & A Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 21. Accessed at http://www.unhcr.org/4cf3ad8f9.html.
[120] B Wigley, *The State of UNHCR's Organization Culture*, UNHCR Evaluation and Policy Analysis Unit, EPAU/2005/08 (May 2005), p. 3.

> background and a support to the achievement of the actual primary purpose of protecting and assisting refugees.[121]

This insight into organizational culture is exemplified by the heavy demands of indicator measurement. It is unclear how the strain of measurement on staff could be lifted short of including in the global budget resources to hire and train independent researchers or reducing the number of indicators on which staff members are required to report.

Indirect consequences of indicator use: innovation and ossification

Indicators, as a tool of knowledge, tend to construct the problems they reveal. The use of quantitative data over time can alter institutional perceptions of the underlying problems that must be confronted. Efforts to measure progress on the basis of a given indicator over time have the potential to deflect attention from wider concerns that are less easily measurable, narrow planned programming, and ossify existing approaches. PDES' evaluation of UNHCR's Results-Based Management roll-out states:

> One of the risks perceived by UNHCR staff in relation to the implementation of RBM is that of focusing on outcomes that are relatively easy to measure, while paying less attention to equally or more important outcomes that require alternative assessment methods.[122]

While data collected through indicators is not intended to cover the field of relevant information (supplementary information is collected via qualitative indicators, disaggregated sectoral data, and narrative reports),[123] the expenditure of time, effort and analysis that goes into the collation of indicators may skew attention towards indicators as the primary source of evidence for decisions on the framing of policy and provision of resources. As the structure of indicators become fixed, it can become increasingly difficult to support budgetary allocations for alternative solutions or to confront new problems.

The potential for these negative indirect effects are inherent in UNHCR's current indicators model. In the area of education, for example, progress on current indicators will require an increase in refugee students enrolled in school programs, rather than ensuring that those currently enrolled are appearing for class or performing at similar levels to their peers.[124] HIV/AIDS outcomes are measured in the *Practical Guide* through "number of condoms distributed per month",[125] the binary question of

---

[121] *Id*, p. 18.

[122] R Allen & A Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 18. Accessed at http://www.unhcr.org/4cf3ad8f9.html.

[123] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. vii.

[124] See e.g *Practical Guide* (2nd ed) (February 2006), Refugee Camp/Settlement Indicators 56-59; Urban Programme Indicators 27-31. This point is also made in R Allen & A Li Rosi, "Measure for Measure: A Field-based snapshot of the implementation of Results Based Management in UNHCR" PDES/2010/13 (1 November 2010), p. 18. Accessed http://www.unhcr.org/4cf3ad8f9.html. It should be noted however that child literacy is the subject of independent indicators.

[125] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), Refugee Camp/Settlement Indicator 38.

whether "stocks of condoms have run out for more than a week",[126] and the question of whether "antiretroviral therapy is available in/for hosting community/refugees".[127] More innovative programs for countering the spread of HIV and AIDS (that are less amenable to quantitative evaluation) are conceivable, such community-based or community-led workshops, the introduction of free and confidential blood tests on a voluntary basis, or the introduction of a comprehensive strategy to counter the problem of 'survival sex' in refugee communities.[128]

Although these strategies could have an impact on the spread of HIV and AIDS in refugee camps and settlement, they would not result in any change in the key indicators on HIV/AIDS currently in place. It is to be hoped that an outcomes-based approach will focus on broader concerns than those fixed for measurement in the current indicators. However the problem of ossification is real and a degree of watchfulness is essential to ensure that programs do not become skewed towards proof of success on the face of the indicator rather than success in countering the underlying problems that the indicator purportedly measures.

Accountability to whom?

An assessment of accountability practices within the UNHCR raises the fundamental question of which individuals or entities are entitled to hold the institution to account. In an influential article on the framing and application of accountability mechanisms, Grant and Keohane posit a basic division between "participation" models of accountability, directed towards those affected by policies, and "delegation" models, responsive to those who entrust powers to the relevant institution.[129]

Although this schema is premised on an assumption that beneficiaries cannot themselves be considered as stakeholders, which potentially negates the image of beneficiaries as rights-holders and promotes a more passive understanding of their role, it is a valuable means for analyzing instruments adopted in diverse accountability models. Making a division between those entrusting an organization with powers and those directly affected by its decisions focuses attention on the range of possible forms of accountability available, and the appropriateness of each in addressing the needs of particular actors.

Applying this schema to the UNHCR, it is arguable that indicators are conceived of as a means of ensuring accountability not only to stakeholders that have entrusted power and resources in the organization, but also to the beneficiaries that are directly affected by UNHCR policies.[130] Stakeholders to whom the organization must answer encompass donors, the host governments on whose territory operations occur, the

[126] *Id*, Refugee Camp/Settlement Indicator 39.
[127] *Id*, Refugee Camp/Settlement Indicator 40.
[128] 'Survival sex', or 'transactional sex' refers to situations where women exchange sex in order to gain food or papers, to avoid violence, or to fulfil material needs of themselves and family members. See E Pittaway, "Making Mainstreaming a Reality – Gender and the UNHCR policy on Refugee Protection and Solutions in Urban Areas – A Refugee Perspective", *Centre for Refugee Research* (19 November 2009), p. 2.
[129] R W Grant and R O Keohane, "Accountability and Abuses of Power in World Politics" (2005) 99 *American Political Science Review* 29, pp. 32.
[130] See UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), pp. 17-18; 33.

General Assembly and the UN Secretary-General. Beneficiaries include internally displaced persons, refugees, persons who have been resettled or refused resettlement, and applicants for refugee status.

The use of indicators is presented as enhancing accountability to stakeholders through the promotion of efficiency and the establishment of a performance-based form of oversight,[131] mechanisms that can be classified under the delegation model.[132] This results-driven model implicitly engages notions of internal organizational efficiency, whereby central offices can monitor the implementation of policies by regional offices.

The sheer number of UNHCR field offices, some of which are in inaccessible locations with less than reliable communications systems, hinders regular contact with the central administration. Indicators are therefore a means to exercise both hierarchical accountability, whereby actions of field offices can be monitored by UNHCR central offices in Geneva and Budapest,[133] and supervisory accountability, whereby the actions undertaken by implementing partners are measured against UNHCR policy objectives.[134]

RBM strategies are relatively responsive to delegation models of accountability. Although it is possible to critique the chosen objectives and techniques of measurement, indicators are performance tools that give feedback on the effectiveness of programming. As regards beneficiaries, however, indicators are considered a tool to bolster a "rights-based approach" to operations and planning.[135] Creating and reporting on indicators is presented as a 'participatory' exercise which builds partnerships and results in "empowerment":

> Persons of concern are no longer mere "beneficiaries", but rights holders who can exercise their rights, including the right to participation.[136]

While the classification of indicators as providing accountability to stakeholders is convincing, the suggestion that they provide participatory accountability to beneficiaries is less compelling. The statement begs the question of exactly how beneficiaries 'participate' in the process of indicator-building, and whether indicators can in truth provide direct accountability to beneficiaries in the form of robust participatory rights.

---

[131] *Id*, p. 33.

[132] See discussion of delegation and participation models of accountability in R W Grant and R O Keohane, "Accountability and Abuses of Power in World Politics" (2005) 99 *American Political Science Review* 29, pp. 29-31.

[133] On hierarchical and supervisory forms of accountability see R W Grant and R O Keohane, "Accountability and Abuses of Power in World Politics" (2005) 99 *American Political Science Review* 29, pp. 36.

[134] This form of accountability is particularly significant given that the growth in UNHCR's functions has been accompanied by greater reliance on implementing partners to fulfill its mandate. Over 960 implementing partners are currently registered with UNHCR. See the UNHCR Partner Directory, www.unhcr.org/cgibin/texis/vtx/search/?page=&comid=4a13fbc46&cid=49aea93a3f&scid=49aea93a24&sort=title (accessed 18 February 2011).

[135] Humanitarian workers routinely argue that the mere creation of indicators is a tool for promoting accountability and participation amongst affected populations. See M L Satterthwaite, "Indicators in Crisis: Rights-based Humanitarian Indicators in Post-Earthquake Haiti" (2011) (forthcoming), p. 11.

[136] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p.17.

Assessing the quality of beneficiary 'participation'

As Arnstein wryly notes, the idea of participation "is a little like eating spinach: no one is against it in principle because it is good for you".[137] Nonetheless, the vocabulary of participation can obscure a range of mechanisms, some of which amount to little more than empty ritual or the laundering of raw power. In order to determine the effect of indicators as a tool of "empowerment", it is useful to establish at what stage of the indicator process the participation of beneficiaries is sanctioned.

Indicators are framed during consultations between UNHCR staff in Geneva and field offices, other agencies in the humanitarian field, and NGO partners.[138] Participation of populations of concern is not envisaged at this point, but rather at the point of *data collection*. Eleven indicators rely on data collected in the specific arena of "participatory assessment" – structured discussions between persons of concern and UNHCR staff aimed at enabling beneficiaries "to participate as partners in the design of programmatic responses to issues shaping their lives".[139]

Urban Programme Indicator 31, which measures the "number of 15-24 year old refugees enrolled in training (non-formal, vocational, skills, etc)" takes this idea one step further by suggesting that "youth and adolescents themselves can be mobilized and trained to collect information on this particular indicator".[140] Another group of indicators notes that "women's or youth associations" can collect information on the literacy rates of adolescent refugees in areas where data is lacking.[141]

The refusal to involve populations of concern at the initial stage of framing an indicator is telling. At the point of framing, political assessments as to which phenomena are the most significant, how they are best proxied, and whether collection can be essentialized across diverse camps remains open. Conversations surrounding framing offer a narrow space in which to shape a process that hardens dramatically as iterative functions take effect – if UNHCR seeks to assess the same phenomena over time, it must continue to measure the same proxy chosen during the initial framing.

A conception of refugees as data-collectors rather than as legitimate participants in planning processes may narrow the space for representation and empowerment. The role lowers the contribution of refugees to a mechanical operation within a pre-conceived system that is not open to challenge or modification.

An element of dependency is implicit in the current scope of refugee involvement in the indicator project. The Listening Project has challenged such conceptions of low-rung participation, stating:

---

[137] S R Arnstein, "A Ladder of Citizen Participation in the USA (1971) 57 *Journal of Royal Town Planning Institute*, 176 at 177. I am indebted to M Mohan, "The Paradox of Victim Centrism", (2009) 9 *International Criminal Law Review* 773 for directing me towards this reference.
[138] UNHCR Division of Operational Services, *Practical Guide* (2nd ed) (February 2006), p. 57.
[139] *Id*, p. 34.
[140] *Id*, Urban Programme Indicator 31. See also Refugee Camp/Settlement Indicator 59 which suggests the same data collection strategy.
[141] *Id*, Urban Programme Indicator 32; Refugee Camp/Settlement Indicator 63.

> Another essential element of participation appears to be the power to decide, instead of only the power to carry out what others have decided.[142]

Successful involvement of persons of concern is not measured by collaboration on an innovative approach to organizing food distribution, or to resolving endemic problems of queues at water pumps and showers, but by their capacity to undertake a head-count of 'percentage of population living within 200 m from water points' (Camp/Settlement Indicator 44), or 'number of persons available per drop-hole communal latrine' (Camp/Settlement Indicator 48).

The anticipated input of refugees devolves in these measurements to a form of assistance rather than critical engagement. Agier's post-colonial critique of UNHCR administration posits that knowledge in camps is only understood through the optic of a 'culture of templates', creating an apparatus of power and profiling that abandons democratic practices.[143] If we are to accept the argument that indicators tend to stymie innovative solutions to problems and instead elevate the production and maintenance of data to a central position in operational practice, the failure to include the subjects of indicators in higher-level dialogues appears doubly problematic.

Looking further afield: are there alternate routes to participatory accountability?

The failure of the current indicator model to provide true participatory accountability is exacerbated by the lack of alternative participatory forums available to UNHCR beneficiaries. The accountability framework in which UNHCR operates is largely self-regulatory. Actions before national courts are hindered by UNHCR's immunity as an international organization,[144] and it is contestable whether UNHCR is subject to human rights principles under international covenants.[145]

In this environment, UNHCR's activities in making final determinations of refugee status and administering territories are subject to three major accountability mechanisms – the Office of Internal Oversight Services (the OIOS), the Inspector General's Office, and the Policy Development and Evaluation Service (PDES) (previously the Evaluation and Policy Analysis Unit).[146] All three of these

---

[142] Collaborative Learning Projects, "The Listening Project Issue Paper: Discuss Together, Decide Together, Work Together (September 2008), p. 3.

[143] M Agier, "Humanity as an Identity and its Political Effects (A Note on Camps and Humanitarian Government)" (2010) 1 *Humanity: An International Journal of Human Rights, Humanitarianism and Development* 29, p. 37.

[144] Although recent decisions in Europe suggest a nascent tendency to lift immunity in cases where an effective remedy is otherwise unavailable: see A Reinisch, "The Immunity of International Organizations and the Jurisdiction of their Administrative Tribunals" (2008) 7 *Chinese Journal of International Law* 285.

[145] For a cogent argument that UNHCR is required to uphold Human Rights standards, see M Pallis, "The Operation of UNHCR's accountability mechanisms" (2005) 37 *NYU Journal of International Law and Politics* 869, pp. 872-875.

[146] For an excellent analysis of the role of the OIOS, the Inspector General's Office, and EPAU (the predecessor to PDES) in providing accountability to UNHCR operations, see M Pallis, "The Operation of UNHCR's accountability mechanisms" (2005) 37 *NYU Journal of International Law and Politics* 869. In addition to these three mechanisms, UNHCR staff members may seek redress for impugned internal administrative decisions via the new UN Dispute Tribunal system. The UN Dispute Tribunal and Appeals Tribunal do not give standing to other stakeholders.

mechanisms – although capable of providing indirect benefits to beneficiaries as a community – are skewed towards providing primary accountability to stakeholders.

The Office of Internal Oversight Services

The OIOS is an independent body established by the United Nations to assist the Secretary-General in fulfilling internal oversight responsibilities throughout the Organization.[147]. Given the scope of the OIOS' activities within the UN it is not practically capable of engaging in 'police patrol' oversight activities, and is generally mobilized after a complaint is made to its confidential investigation hotline.[148] Reflecting on the fact that both of the major OIOS investigations into UNHCR operations have been requested by UNHCR itself, Pallis states:

> On one hand, this may be a testament to UNHCR's willingness to open itself to scrutiny… But on the other hand, as far as UNHCR is concerned, the OIOS is not a "watchdog" on the lookout for potential problems, but rather a standing facility for ad hoc investigation.[149]

In practice, OIOS is more analogous to a self-regulatory mechanism than to a robust oversight body. As such it is a useful mechanism for ensuring internal accountability but is not equipped to directly engage with UNHCR beneficiaries.

The Inspector-General's Office

The Inspector-General's Office is charged specifically with evaluating and investigating UNHCR's management practices. In particular, the Office is responsible for assessing the quality of UNHCR's management, investigating misconduct allegations when they are raised,[150] and undertaking inquiries into attacks on UNHCR personnel and incidents that potentially affect the Organization's reputation.[151]

At first glance the Office appears to provide a stronger avenue for direct claims by persons of concern than the OIOS. Allegations of misconduct listed on the Inspector-General's Office online complaint form include *inter alia* sexual exploitation or abuse, fraud (via corruption and bribery or in relation to decisions on Refugee Status Determinations), and abuse of authority.[152] However the structure of the Office is geared towards receiving complaints made by UNHCR staff rather than those made by persons of concern.[153]

The Office is based solely in Geneva, and does not operate a toll-free telephone service, hindering the ability of refugees in camp situations to access the Office.[154] Further, the Office treats complainants not as *victims* of misconduct in pursuit of

[147] The OIOS is divided into three main divisions; the internal audit division, the investigations division, and the inspection and evaluation division. It is authorized to engage in monitoring, internal auditing, inspections, evaluations, and to report on matters falling within its mandate. For further information on the establishment and role of the OIOS see http://www.un.org/depts/oios/ (accessed 1 January 2011).

[148] This term is adopted from the model in K Raustiala, "Police Patrols & Fire Alarms in the NAAEC" (2004) 26 *Loyola Los Angeles International and Comparative Law Review* 389, pp. 390-391.

[149] M Pallis, "The Operation of UNHCR's accountability mechanisms" (2005) 37 *NYU Journal of International Law and Politics* 869, p. 891.

remedies, but rather as *reporters*, or 'fire alarm' triggers that assist in identifying and rooting out general incidences of misconduct.[155]

A factsheet prepared by the Inspector General's Office clearly states that a complainant to the Office is not entitled to be updated on the progress of an investigation and will not be informed of its outcome.[156] As such the body can be considered as an important evaluative tool for stakeholders in its aim of supporting "effective, efficient and accountable management of UNHCR operations",[157] but a less than effective tool for directly responding to the needs of beneficiaries.

The Policy Development and Evaluation Service

The Policy Development and Evaluation Service (PDES), which replaced the EPAU in 2006, is a UNHCR think-tank that reports to the High Commissioner and is committed to analyzing and publishing evaluative reports on UNHCR "policies, programmes, and practices".[158] The office has four core international staff members and one administrative staff member, and receives an annual consultancy budget of $300,000.[159]

While PDES, like its predecessor EPAU has proved a balanced critic of UNHCR management shortfalls, it likewise cannot be considered to encompass a 'participation' model of accountability under which beneficiaries are entitled to hold UNHCR to account.

---

[150] United Nations Staff Rule 110.1 defines misconduct as "*failure by a staff member to comply with his or her obligations under the Charter of the United Nations, the Staff Regulations and Staff Rules or other administrative issuances, or to observe the standards of conduct expected of an international civil servant*". See UNHCR Inter-Office Memorandum No. 054/2005, "The role, functions and *modus operandi* of the Inspector General's Office" (3 November 2005) at para. 5.2.1. Accessed at www.un.org/en/pseataskforce/docs/memo_on_the_role_and_function_of_the_inspector_generals_offi.pdf.

[151] See IGO website at http://www.unhcr.org/pages/49f0619f6.html (accessed 29 December 2010).

[152] IGO Online Complaints Form, accessed at www.unhcr.org/cgi-bin/texis/vtx/page/complaintsigo.html.

[153] It should be noted however that complaints by beneficiaries are anticipated in policy documents. See e.g. UNHCR Inter-Office Memorandum No. 054/2005, "The role, functions and *modus operandi* of the Inspector General's Office" (3 November 2005), para. 5.11.3. Accessed at http://www.un.org/en/pseataskforce/docs/memo_on_the_role_and_function_of_the_inspector_generalsoffi.pdf.

[154] See the means of contact outlined on the IGO website at http://www.unhcr.org/pages/49f064786.html.

[155] See *supra*, ft 148.

[156] UNHCR Factsheet, *How to Report Misconduct and What to Expect*, Accessed at http://www.unhcr.org/4a1278f06.pdf.

[157] See UNHCR Inter-Office Memorandum No. 054/2005, "The role, functions and *modus operandi* of the Inspector General's Office" (3 November 2005), para. 2.2. Accessed at http://www.un.org/en/pseataskforce/docs/memo_on_the_role_and_function_of_the_inspector_generalsoffi.pdf.

[158] See PDES overview at http://www.unhcr.org/pages/4a1d28526.html.

[159] "PDES briefing to UNHCR Executive Committee", *Informal Consultative Meetings* (3 Feb 2009). Accessed at http://www.unhcr.org/4987043b2.pdf.

Participation in practice

An overview of the accountability mechanisms currently in place in UNHCR operations reveals a strong bias in favour of delegation mechanisms aimed at satisfying stakeholders over participation models aimed towards beneficiaries themselves. The lack of alternate mechanisms sharpens the need to engage beneficiaries directly in the standards and indicators project, given its increasing centrality in organizational planning and aid distribution. The current framework of using participatory assessments as a data source for measurement of pre-determined aims is an inadequate method of engaging beneficiaries in the planning process.

There are legitimate practical questions as to how participation could be achieved at an earlier phase of framing. The diversity of languages, ethnic groupings, experiences and locations of UNHCR beneficiaries will render any attempt to create a representative sample of voices difficult. Nonetheless, the very complexity of refugee experiences attests to the significance of maintaining a plurality of opinions in the room during initial framing discussions. The creation of indicators by its nature universalizes concerns and monitoring practices throughout UNHCR operations. Input from individuals subject to such operations is a crucial aspect of nuanced and appropriate framing.

Rather than conceiving of refugee, asylum seeker or IDP participants as a mouthpiece for a particular group or community, it would be preferable to encourage participants to act in their individual capacity.[160] Several methods of choosing participants are feasible. A lottery for interested parties or a voting system within a range of individual camp environments (including IDP camps, long-term refugee camps and self installed camps under surveillance) is a conceivable, if administratively burdensome, method of ensuring equitable access to discussions. Linguistic concerns could be alleviated by the presence of qualified interpreters chosen after the election or selection of participants. Such approaches would provide practical – if preliminary – steps towards enhancing the participatory role of persons of concern from a purely mechanical function to that of active members of the decision-making process.

Conclusion

UNHCR seems unlikely to retreat from its growing reliance on indicators in the short-term. Although framing may be improved and tinkering with terms will continue,[161] internally generated standards and indicators appear rooted as a significant aspect of governance in UNHCR operations. The breadth of UNHCR's functions in the global sphere enhances the need for analytic clarity in the construction and use of these indicators, and robust discussion of their shortcomings. It is hoped that a spirited debate on the role and limits of indicators within UNHCR acts to hone indicators'

[160] It is accepted that an approach to choosing participants that remains entirely neutral, to the extent of choosing ten male participants and no female participants, would be flawed. To be guided by an effort to maintain a plurality of experience does not, however mandate that individuals selected must negate their own ideas in order to represent the broad interests of their gender, ethnicity or social group.

[161] As recently as December 2010, UNHCR management declared an intention to sharpen indicators under the *Results Framework* where required: see *Measure for Measure: A Field-Based Snapshot of the Implementation of Results Based Management in UNHCR: Initial Management Response* (1 December 2010), p. 3. Accessed at http://www.unhcr.org/4d1b47d49.html.

structure and methodology, assists in providing adequate support and training to data collectors in the field, and resolves issues of participatory accountability. Closer attention to the pitfalls of data collection and the capacity of indicators to obscure contested policies may also assist in instilling a more critical attitude towards indicators – one that is alive to the limits of their comparative function, and their potential to narrow solutions to those conducive to quantitative measurement rather than focusing attention on those most capable of providing practical results.