

Transitional Justice Review

Volume 1 | Issue 4

Article 9

March 2016

Notes from the Field: Global Indicators for Transitional Justice and Challenges in Measurement for Policy Actors

Paige Arthur

Public Action Research, parthur@publicactionresearch.com

Follow this and additional works at: <https://ir.lib.uwo.ca/tjreview>

<http://dx.doi.org/10.5206/tjr.2016.1.4.9>

Recommended Citation

Arthur, Paige (2016) "Notes from the Field: Global Indicators for Transitional Justice and Challenges in Measurement for Policy Actors," *Transitional Justice Review*: Vol. 1 : Iss. 4 , Article 9.

DOI: <http://dx.doi.org/10.5206/tjr.2016.1.4.9>

Available at: <https://ir.lib.uwo.ca/tjreview/vol1/iss4/9>

This Article is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Transitional Justice Review by an authorized editor of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.

**Notes from the Field:
Global Indicators for Transitional Justice and Challenges in
Measurement for Policy Actors¹**

Paige Arthur, Public Action Research
parthur@publicactionresearch.com

Abstract

Indicators have become an important tool for policy actors at the bilateral and multilateral level over the past twenty years; however, they have mainly been developed in relation to development and public health goals. This note identifies the practical and methodological challenges in developing global (i.e. cross-national) indicators for transitional justice, through reflection on a practical engagement with UN Women, for which the author developed two indicators on women's participation in truth commissions and in reparations programs. Specific challenges to developing the indicators included: the lack of administrative data on transitional justice; difficulty in establishing agreed definitions on "what" is being measured, which is linked to the lack of common agreement on the objectives of transitional justice initiatives; lack of standardization of data collection practices across countries; lack of engagement between transitional justice institutions' staff and statisticians; and the general challenges in measuring progress against human rights objectives. I introduce a "basket" approach as an imperfect solution to this data reality. The

¹ I thank the following people for their substantive input on the ideas in this essay: Nahla Valji, Alison Davidian, Sara Duarto Velero, Ruben Carranza, Cristián Correa, Elena Naughton, Peter van der Auweraert, and Sofia Macher. The approach described in the essay is, however, my own.

note concludes by identifying specific changes that would ease the process of developing meaningful cross-national indicators on transitional justice.

Introduction

Over the past several decades, development policy actors in the World Bank and the United Nations have led the way in creating indicators that permit cross-national comparisons of countries' performance across a variety of development-related dimensions.² These indicators rely largely on administrative and statistical data produced by states themselves, often through their national statistical bodies. Human rights policy actors have lagged behind, owing both to a lack of similar data and, relatedly, lack of international consensus on "what" should be measured.³

Like the broader human rights field, the TJ field lacks both data and agreement on "what" should be measured in cross-national indicators for policy actors. In this "note from the field," I will discuss one novel attempt to develop global, cross-national indicators related to TJ. I define "indicators" as signals of change, usually quantitatively expressed.⁴ Such indicators are different from the data sets that

² See, for example, the UN Development Programme's (UNDP) *Human Development Reports* and the Millennium Development Goals; as well as the World Bank's World Development Indicators.

³ Unlike in the development field, statistics can be difficult to come by, as states often cannot be trusted to gather accurate data on their own human rights records. While data from independent agencies or organizations may appear to fill this gap, these data may not adhere to a recognized, agreed set of standards and methods; moreover, states may criticize these agencies (fairly or not) for lack of impartiality. Finally, human rights remain a deeply contested issue for many states, some of which may wish to prioritize economic and social rights over civil and political rights.

⁴ A more technical definition: "An indicator is *a collection of named, rank-ordered, simplified and processed data that purports to represent the past or projected performance of different units*. An indicator simplifies and processes data about a named social phenomenon in a way that makes it possible to compare and evaluate units such as countries, communities, organizations, or individuals" (italics in original). Kevin E. Davis and

scholars have been producing over the last half-decade or so, in that their intent is to capture countries' commitments to and performance on TJ.

In 2013, I worked with UN Women to develop two global indicators on TJ, as part of a larger framework designed to monitor the implementation of the Security Council's landmark Resolution 1325 on women, peace, and security (SCR 1325).⁵ As this note will show, there are significant obstacles to developing such indicators, and all of them are important lessons that can help inform the practice of both national and international actors working on TJ—not just international policy actors like UN Women staff. First, little data is produced on TJ initiatives (by TJ actors, states, multilateral actors, etc.), and what data there is may not be accessible to outside parties. Second, whatever data currently exists may not be particularly relevant to the decision making needs of a range of actors, including the global policy actors in UN agencies like UN Women. Third, there is no agreement internationally on what should be measured; even at the national level, there is little to no discussion of this issue. Finally, there is a need to develop indicators across a range of dimensions—just like development actors currently do—in order to make judgments whether TJ initiatives are performing well or poorly. In the case I discuss in this essay, I proposed a “basket” of several indicators, rather than a single indicator. While the basket approach is not perfect, it is more relevant and useful than a single indicator; this approach has the advantage of capturing five different dimensions of a reparations performance regarding the effective inclusion of women and girls.

To explore how these lessons were arrived at, this note will start with some background to the indicators project, including the reason why UN Women felt it needed the proposed indicators. It will

Benedict Kingsbury, *Indicators as Interventions: Pitfalls and Prospects in Supporting Development Initiatives* (New York, NY: Rockefeller Foundation, 2011): ii.

⁵ Security Council Resolution 1325 (2000); available from <http://daccess-dds-ny.un.org/doc/UNDOC/GEN/N00/720/18/PDF/N0072018.pdf?OpenElement>.

then move to a discussion of my approach to measuring the indicators, including a description of the normative basis on which the indicator was based. This normative element is important because it shows the range of dimensions that any indicator of TJ might need to capture. Next, I turn to the challenges with both the conceptual relevance and measurement of the proposed indicators. The challenges suggest that the need for an alternative approach—the basket approach—which I then describe in detail. The essay concludes with some reflections and lessons learned.

Background

SCR 1325, adopted in 2000, was intended to increase the participation of women and incorporate a gender lens into all of the UN's peace and security efforts. It has been an important tool for actors both inside and outside of the UN to promote inclusion of women and girls—but the question remains, how effective has that tool been, and what, if anything, is changing as a result? These are the questions that motivated the UN, on the tenth anniversary of SCR 1325, to develop a strategic framework and an inter-agency task force to monitor SCR 1325 on a forward-going basis.⁶ Of the many indicators developed for the strategic framework, I worked on the only two that directly reference TJ initiatives:

- Indicator 25: Extent to which Truth and Reconciliation Commissions include provisions to address the rights and participation of women and girls
- Indicator 26a: Percentage of benefits from reparations programmes received by women and girls

⁶ See UN Strategic Results Framework on Women, Peace and Security, 2011–2020; available from http://www.un.org/womenwatch/ianwge/taskforces/wps/Strategic_Framework_2011-2020.pdf.

I was asked to develop a methodology to measure each of these, working closely with the lead agency, UN Women.⁷ The end result was a nearly 20,000-word paper, as well as two “tools” that UN staff could use to put the indicators into practice. The hope was that data could be generated in advance of the next report of the Secretary-General on Women, Peace and Security in 2014.⁸

In the end, however, it is unclear whether the data for the indicators have actually been developed (although it is said that the paper has spurred useful dialogue on how to approach the issues within the UN). The reasons for this tenuous outcome should be understood in light of the conceptual and practical difficulties I faced in collaboration with my UN Women colleagues in this unexpectedly ambitious task.

Conceptual Relevance: Why this Indicator?

In this note from the field, I will focus on Indicator 26a, “Percentage of benefits from reparations programmes received by women and girls.”⁹

⁷ Specifically, I was asked to present a proposal for operationalizing both indicators, including (i) refinement of the parameters of each indicator and what they should measure, (ii) methodologies for data collection against these defined indicators, (iii) countries where the indicators should be implemented, and (iv) recommendations for future refinement/revision of both indicators under broader UN policy processes.

⁸ For reference to the indicators developed in this paper, see “Report of the Secretary-General on women, peace and security” (S/2014/693), 25; available from http://www.securitycouncilreport.org/atf/cf/%7B65BFCF9B-6D27-4E9C-8CD3-CF6E4FF96FF9%7D/s_2014_693.pdf.

⁹ Because this indicator asks for a proportion of benefits, it has a clear quantitative element for which—on the face of it—it might seem easy to develop a methodology. By contrast, the other indicator asks to assess “the extent” to which truth commissions include provisions to include women and girls. For this indicator, I developed a scorecard approach that would allow an independent scorer (i.e., UN staff member) to rate various aspects of a truth commission’s design and implementation across a range of criteria cited as important in key documents on

The motivation for gathering data on this issue stemmed from the recognition that reparations programs have not had a positive track record in ensuring that women and girls benefit from their initiatives and, indeed, that many have implicitly discriminated against women.¹⁰ Key gaps include the fact that, in many cases, sexual and gender-based violence (SGBV) has not been treated comprehensively as a reparable harm. In particular, few programs have included repair for forced union, forced pregnancy, forced sterilization or loss of fertility, and forced abortions. Even in cases where SGBV is included as a reparable harm, women are less likely to come forward to claim benefits, often owing to the cultural stigma of sexual violence.¹¹ Moreover, socio-economic violations such as displacement and loss of property, which women suffer disproportionately in times of conflict, are usually not included as reparable harms.

Additionally, compensation schemes may not adequately assess logistical challenges for including women, for example, the fact that women are less likely than men to have identification, proper official documents/certificates, or bank accounts; for these reasons, money intended for them may be directed to male relatives, which can cause family conflict.¹² Compensation schemes sometimes also reproduce local inheritance customs that discriminate against women.

gender and TJ. This more subjective approach was appropriate for the intent and wording of this indicator.

¹⁰ The examples that follow are drawn from: UN Women, “A Window of Opportunity: Making Transitional Justice Work for Women,” 2nd ed. (2012), 16–17, available from www.unwomen.org/wp-content/uploads/2012/10/06B-Making-Transitional-Justice-Work-for-Women.pdf and Rubio-Marín, *What Happened to the Women*.

¹¹ This was the case in South Africa. See Beth Goldblatt, “Evaluating the Gender Content of Reparations: Lessons from South Africa” in Ruth Rubio-Marín, ed., *What Happened to the Women: Gender and Reparations for Human Rights Violations* (New York: SSR, 2006), 69.

¹² Other hurdles include illiteracy and linguistic barriers. For a discussion, see Rubio-Marín, “Introduction” in *What Happened to the Women*, 34; and for the South African example, see Goldblatt, “Evaluating the Gender Content of Reparations,” 73–74.

Finally, reparations may overlook the indirect impact of human rights violations—for example, how the loss of a male family member can have strongly negative consequences for women and children. Additionally, women may not be adequately consulted on decisions about how to create collective reparations programs.

Indeed, discussions with UN Women program staff revealed that **all** of these issues were important to them, and that ideally an indicator could be developed to capture as many of these dimensions as possible. This desire, although completely understandable from a policy standpoint, suggested that program staff actually required a larger set of indicators, which was not realistic given the scope of the project. Thus, a single statistic was not what program staff needed to make good judgments about country-level performance; they likely would have found a scorecard or index approach more useful.

This reaction from program staff can be contrasted with that of the statistician from the Evaluation Unit, who strongly emphasized the need for a single statistic that could be framed in such a way as to be easily comparable across countries. One suggestion she made was to develop an indicator that showed spending on reparations for women/girls as a proportion of GDP. This approach would probably be useful from a development perspective, but it was not particularly useful from a human rights perspective. This difference in approaches was striking throughout the project. Indeed, I discovered that the Office of the High Commissioner for Human Rights (OHCHR) has developed its own manual on developing human rights indicators, which suggests that output and/or outcome indicators—rather than financial/budgetary indicators—are preferable in assessing the enjoyment of rights.¹³ This publication was unknown to the statistician from the Evaluation Unit, however, and it is unclear how widespread (if at all) this kind of approach is.

¹³ OHCHR, *Human Rights Indicators: A Guide to Measurement and Implementation* (Geneva, 2013), 79.

Balancing these different perspectives—each legitimate in its own right—was an important part of the project.

Approach

My approach relied on document review, including key normative and other background documents on TJ developed primarily by the UN, and interviews with technical and country experts. On the basis of these documents and expert interviews, I clarified the underlying objective for each indicator, developed definitions of terms, and set forth criteria for assessing change over time. I then created tools and guidance for collecting data on each indicator, based on best practices and similar examples within the UN system. Finally, I reviewed these criteria against several case studies—relying on document review and interviews with experts—to assess the limitations of the proposed methodology and to make changes to the tools and guidance as needed. In the case of the reparations indicator, I tested the proposed methodology in the case of Peru, to see if it would realistically show change over time.

Another important aspect of my approach was maintaining a consistent dialogue with my counterparts at UN Women. Two of these were program staff with expert knowledge on TJ, and one was a statistician working in the Evaluation Unit, with a strong background in development, but little familiarity with TJ.

Finally, because the indicators were being developed for UN use, they had to follow norms as laid out in key UN documents. This was useful, in the sense that key definitions (e.g., “reparations program”) could just be adopted from this guidance. It was less useful, however, because the guidance itself is complex—and this complexity needed to be reflected in the indicators. Thus, although the most obvious form that reparations can take is cash payment, reparations

programs should offer a variety of forms of repair, according to the *Basic Principles and Guidelines*.¹⁴ These are:

¹⁴ UN General Assembly resolution, “Basic Principles and Guidelines on the Right to a Remedy and Reparation for Victims of Gross Violations of International Human Rights Law and Serious Violations of International Humanitarian Law” (2005), Arts. 19–23, available from www.un.org/ga/search/view_doc.asp?symbol=A/RES/60/147.

- **Compensation:** for example, cash payments or pensions
- **Restitution:** for example, of lost property, civic status, or jobs
- **Rehabilitation:** for example, of health or mental well being
- **Satisfaction:** for example, public apologies and full disclosure of the facts about state-led abuse
- **Guarantees of non-repetition:** for example, reforms of abusive state institutions

While the normative guidance recognizes that it is impossible to fully repair the harms of gross human rights violations, programs that have taken a more complex approach to their mix of benefits have perhaps been more successful in achieving basic TJ goals such as recognizing victims and fostering a sense of citizenship—for example, programs in Chile and Argentina that provide a mix of payments, social benefits, and symbolic actions like memorials and apologies. With respect to recognizing victims, many programs have done this successfully, especially through the integration of individualized symbolic gestures with material benefits like cash payments and pensions—for example, including personalized letters of apology or a copy of the truth commission’s report when victims are sent their reparations check. The need for a comprehensive and complex approach to reparations was integrated into the conceptualization of the indicator, as I demonstrate below.

Key Challenges

Although the indicator seems simple and straightforward, in fact it posed a variety of difficult challenges, most importantly with respect to the conceptual relevance of the indicator and the ease of actually measuring, or producing data, for the indicator.

Problems with Conceptual Relevance

The most important challenge relates to conceptual relevance, meaning the indicator's relevance to the change that it is intended to signal. In fact, Indicator 26a, as written, would not provide a reliable signal of the specific change that the UN seeks, which is the effective inclusion of women and girls in reparations programs. Instead, this indicator would show positive and negative fluctuations over time that may have little to do with how well a program has actually included women and girls.¹⁵

For example, in Sierra Leone, women and girl victims of sexual violence were offered fistula surgery as part of the reparations program. This is something that, once completed, does not need to be offered again. The result, then, is that the total share of reparations benefits for women and girls would see a **negative** change after most women had received their fistula surgeries, even though women may still be well served by the reparations program as a whole. These kinds of positive and negative fluctuations will be seen in every reparations program that offers one-off benefits that are enjoyed primarily by male or by female beneficiaries. Ultimately, I ran a series of models to substantiate this conclusion—i.e., that the indicator, as constructed, would not necessarily change over time in a way that could easily be construed as a “good” or “bad” outcome in terms of the change desired.

Moreover, the indicator as currently constructed would not say anything about changes in one of the key obstacles to women and girls' inclusion that the UN has identified, which is the fact that many reparations programs do not include SGBV abuses as reparable harms. Discussions with UN Women program staff suggested that this was a critical signal that the indicator needed to make. However, in theory, a

¹⁵ Although I did not have the opportunity to address this in the paper, I would also add that reparations programs ought not to treat “women” and “girls” as a single group; there is an emerging body of research on the inclusion of children in TJ that discusses some of the differences. See ICTJ's “Children and Youth” program page at www.ictj.org/our-work/transitional-justice-issues/children-and-youth.

reparations program could see a high percentage of benefits going to women and girls even when SGBV abuses are not being repaired—for example, if benefits are given to widows of men who have been killed or disappeared.

Problems with Measurement

Other significant challenges relate to methodological and measurement issues specific to reparations programs. First, reparations benefits are quite different in kind, not just within countries, but also across countries. Although there was an expectation (mainly from the statistician from the Evaluation Unit) that I could create a complex formula to aggregate different aspects of a reparations “program,” as is done with other types of social programs, this is not realistic or even desirable. This would assume that such programs are similar across countries (they are not), or that they have a “program-like” quality, which many do not. In fact, it is not always easy to identify what counts as a reparations program.¹⁶ While some countries, such as Brazil, Malawi, and Morocco, have established self-standing reparations commissions to administer benefits, other countries have not. For example, in Argentina, reparations have been legislated incrementally through discrete laws, and no particular administrative body is in charge of them.¹⁷

As an in-depth example, consider Nepal’s Interim Relief Program (IRP), which provides benefits to families of people murdered or disappeared during the conflict.¹⁸ The program provides

¹⁶ In fact, OHCHR suggests that most countries do not have reparations “programs,” which it defines as “designed from the outset as a systematically interlinked set of reparations measures.” Instead, it observes that most countries have reparations “efforts” that emerge in an ad hoc, incremental way. OHCHR, “Rule of Law Tools for Post-Conflict States: Reparations Programmes” (2008), 3, available from www.ohchr.org/Documents/Publications/ReparationsProgrammes.pdf.

¹⁷ *Ibid.*, 12.

¹⁸ Information on Nepal’s IRP comes from Ruben Carranza, “Relief, Reparations, and the Root Causes of Conflict in Nepal” (ICTJ, 2012), as well as discussions with

a mix of benefits: one-off cash payments to the nearest relative of the murdered/disappeared; additional one-off cash payments specifically to widows/wives of the murdered/disappeared; and educational scholarships for up to three children paid to the nearest relative of the murdered/disappeared. Further, health benefits are provided to persons who were injured during the conflict. These benefits entail free medical services from government-run agencies, or reimbursement of health services from private doctors.

How could all of this be combined into a single percentage? While it might seem straightforward to calculate the percentage of the one-off cash payments received by women/girls, there are some ambiguities. For example, in cases where children are the “nearest relatives,” then the payments will go to an adult who is a guardian, who may be a male relative. So who has actually “received” the benefits in this case—the male guardian or the children, including girls?¹⁹ Other parts of the program are also ambiguous. First, we do not know if or how many girls benefit from the educational scholarships, because they are limited to three per family, and therefore in larger families the benefits could be used for three boys. Moreover, most health benefits are provided as part of the larger national health services budget, rather than as a separate program—therefore, data cannot be disaggregated so that we know who received what services and how much they cost, except perhaps in the rare cases that people ask to be reimbursed for services from private doctors (typically, this only takes place when the victim is outside of Nepal and cannot access the government-run system). These ambiguities exist in almost every reparations program.

Ruben Carranza and Elena Naughton of ICTJ. See also a previous ICTJ report, “From Relief to Reparations: Listening to the Voices of Victims” (ICTJ, 2011).

¹⁹ Although one could argue that all children, including girls, benefit even if a male custodian receives reparations payments, in South Africa, attempts were made to direct payments to female custodians of children, on the assumption that they were more likely than men to use the money to benefit the child. See Goldblatt, “Evaluating the Gender Content of Reparations,” 73–74.

A second, and perhaps more decisive, measurement challenge concerns the availability of administrative data. This is a general challenge with all TJ measures, not just reparations programs. But reparations present their own special difficulties. It is unrealistic to expect that we will see clear budget lines for each type of reparations benefit. In most countries, the data does not exist in a central location, as programs are often administered by different local and national state agencies. Taking the example of Nepal's IRP once again, there is little accessible data on the costs for health benefits. Eligible beneficiaries go to their local clinics for treatment, and the costs for these benefits are rolled into the state's general health budget—they are not disaggregated anywhere as costs "for victims." While in theory it may be possible to get the data for victims who have used private medical services, and who thus must submit receipts for reimbursement, these instances are rare, and it is unknown whether data is kept on the gender of those who use private medical services. There is a similar problem with the data for educational scholarships. There simply is no data on how much these benefits actually support the schooling of girls, especially since, as mentioned, the scholarships are limited to three children per family, and no data is kept on which children in each family actually use the benefits. Finally, when there is more than one type of benefit, we usually have no way of knowing who or how many people are receiving multiple benefits, and who or how many people (including women/girls) are not. There is not a central ledger anywhere, in which benefits distributed are tracked against particular victims' names.

A Way Forward?

Given the conceptual and measurement challenges discussed, both the program and evaluation staff at UN Women agreed that I should propose an alternative approach.

A number of alternatives were explored, most of which were quantitative in nature and relied on administrative data. UN Women emphasized that the indicator should be easy to measure, given

available data, which placed limits on the kinds of indicators that could be considered. In the end, full methodologies were developed for two other quantitative indicators, before they were rejected. Both of these focused on **counting beneficiaries/victims, rather than counting benefits received**—simply because we have better data on the former than on the latter.

The first proposal was: “Percentage of beneficiaries listed in a reparations programme’s official registry of victims who are women and girls.” This indicator would have relied on existing administrative data in the official registry of victims. While these data are usually readily available, this option was rejected because its conceptual relevance and utility to UN Women staff was unclear. There was, in particular, an objection to a narrow focus on potential beneficiaries rather than actual beneficiaries²⁰—UN Women program staff (understandably) wanted the indicator to signal the degree to which women/girls had actually benefitted from such programs.

The second was: “Percentage of actual beneficiaries of reparations programmes who are women and girls.” This indicator would have relied on administrative data from a range of ministries responsible for aspects of the reparations program. While this approach had the merit of counting “actual” rather than “potential” beneficiaries, the conceptual relevance and utility for UN Women staff was still unclear, specifically because it did not provide any signal of whether harms related to SGBV were being repaired or not. Also, in some cases, the data would not be centralized and might be difficult to generate.

Other possible indicators were discussed, without the development of methodologies. These were also based on data that

²⁰ Unfortunately, the mere act of being registered as a victim in a registry does not mean that a person will actually receive reparations. This has been particularly true in the case of women and girls, for some of the reasons mentioned above (lack of bank accounts, receipt by male relatives or guardians, etc.). There is also the fact that reparations benefits are sometimes promised and then never delivered.

was more readily available, for example, data on one-off payments rather than health benefits.

For example, we discussed the indicator, “Ratio of the percentage of one-off benefits received by women/girls to one-off benefits received by men/boys.” This indicator would have revealed differences in the rate that women/girls received benefits, if there were such differences. However, it was considered to be too one-dimensional. Additionally, it was said that using a ratio would be problematic, as it would be difficult to know whether overall changes were a result of changes in the numerator or the denominator.

We also discussed, “Proportion of estimated women/girl victims who actually register for reparations.” This indicator would be useful in helping to assess the degree to which women are coming forward to register as victims/beneficiaries, and the ease with which they are recognized as victims/beneficiaries. The challenge, however, is that in many cases there are no existing population surveys of estimated victims to make the comparison against.

A “Basket” Approach

Indicators exist in order to help us measure change and to compare similar phenomena across countries. They also help to set policy priorities and to benchmark or measure performance. Human rights indicators, as pointed out in OHCHR’s *Human Rights Indicators*, should be anchored in the normative content of specific rights, and they should focus on measuring the commitments of duty bearers, primarily states.²¹

In this case, **an indicator is needed to signal changes in states’ commitments to the effective inclusion of women and girls in reparations programs.** The challenge is to find an indicator that can capture various attributes of inclusion.

²¹ OHCHR, *Human Rights Indicators: A Guide to Measurement and Implementation* (Geneva, 2013), 33.

Given the complexity and variety of reparations programs, these various attributes related to state behavior and normative change can be captured more easily if a “basket” of relevant indicators is created.²² The proposed title for this revised indicator is: **Women and girls’ effective inclusion in reparations programmes.**

The basket itself would consist of the following indicators, each of which signals one aspect of gender sensitivity for each of the five types of reparations outlined in the *Basic Principles and Guidelines*. Each indicator should be assessed a score of “2” if it has been fulfilled, “1” if it has been partially fulfilled (that is, if it has begun and is still in progress), and a score of “0” if it has not been fulfilled at all. That is, the difference between a “1” and a “2” score is the difference between an ongoing and a completed activity. A short, one-paragraph description should be included to give the rationale for the score for each of the five indicators (see Appendix for more detail).

- **Compensation: The reparations programme has implemented cash payments and/or pensions specifically for widows/spouses.** One key challenge in terms of compensation is the fact that women may be overlooked as “indirect” victims of human rights violations; indeed, the loss of a male head of household can have especially dire consequences for women and families. Experience has shown that if there is not specific language to prioritize widows/spouses in reparations programs’ laws or guidelines, then male family members of the deceased/disappeared may step forward to claim benefits instead of wives/widows. This indicator requires the use of specific language to include widows/spouses in laws and/or implementing guidelines for reparations programs. If widows/spouses are not specifically

²² This approach is modeled on baskets of indicators proposed in DPKO and OHCHR’s publication, *The United Nations Rule of Law Indicators: Implementation Guide and Project Tools* (2011), sec. 3.2; available from www.un.org/en/peacekeeping/publications/un_rule_of_law_indicators.pdf.

mentioned, but only implied, then this should be given a “0” score, even if the cash payments in general have been partially or entirely completed.²³

Restitution: A restitution programme exists, and it receives a score of “2a” or above according to IASC Gender Marker codes (i.e., the project is designed to contribute significantly to gender equality). During conflict, women and girls often suffer disproportionately from socio-economic violations such as displacement and loss of property. Many reparations programs do not, however, address such harms. It will therefore be important to track the inclusion of restitution initiatives in relation to property, civic status, jobs, etc., as part of reparations programs—as well as steps states are making to ensure the effective participation of women and girls in those initiatives. Gender marker codes have been in use by many UN agencies for many years. IASC’s version of these codes has a well-developed set of generic criteria that are readily applicable to restitution programs.²⁴ To determine the proper gender marker code, the IASC criteria must be referenced.²⁵ A program can be coded on a scale of 0 to 2. A score of “0” represents “no visible potential to

²³ Cristián Correa has noted that in Argentina’s reparations program for the disappeared, the “beneficiary” of the payment is the disappeared person him- or herself, rather than a spouse/widow and other family members. As heirs to the beneficiary, widows/spouses do however receive reparations payments. In this kind of ambiguous case, what is most important is that widow/spouses are specifically foreseen to benefit, even if they are not technically “beneficiaries” according to the law.

²⁴ For an overview of the codes, see Global Protection Cluster, “IASC Gender Marker – Frequently Asked Questions,” July 29, 2011; available from www.globalprotectioncluster.org/_assets/files/partners/Gender_Marker_FAQ_29_July2011_EN.pdf.

²⁵ The document can be found on the Global Protection Cluster web site, at www.globalprotectioncluster.org/_assets/files/partners/Gender_Marker_FAQ_29_July2011_EN.pdf.

contribute to gender equality.” A score of “1” represents the “potential to contribute in some limited way to gender equality.” The code of “2” is split into “2a” and “2b.” Both signal programs that mainstream gender into analysis and design. However, “2a” sets a slightly lower bar, in which programs must simply have the “potential to contribute to gender equality,” whereas programs coded as “2b” must have gender equality “as their primary goal.”²⁶ In order to code the restitution program, all aspects of its design and implementation should be considered. Specific signals that the restitution program is designed to contribute significantly to gender equality can include, but are not limited to: wives, including common law wives, are given an equal share of the property being restituted to their husbands if they were married at the moment of the dispossession; widows or wives of the disappeared have immediate access to claim salaries, labor compensation, and property of their husbands free of fees and taxes; and there is an expedient method for wives of the disappeared to declare the absence due to the disappearance of their husbands and to claim inheritance or administer their estates.²⁷

· **Rehabilitation: The reparations programme has implemented medical services specifically for victims of SGBV.** Another key challenge with reparations programs has been the exclusion of SGBV as a reparable harm—especially for forced union, forced pregnancy, forced sterilization or loss of fertility, and forced abortions. In countries where most of the victims of human rights abuse have been male detainees, some states have assumed that SGBV is not relevant, even though we know that men routinely suffer from SGBV and may report it at an even lower rate than women. This indicator

²⁶ Ibid.

²⁷ I thank Cristián Correa for these examples.

will provide a signal not only of whether the need for rehabilitation is being addressed, but also whether or not a reparations program includes SGBV. In order to code this indicator as a “2” or a “1,” the program must include specialized medical services for SGBV victims. Examples of such services include, but are not limited to: psychosocial support; pre- and post-natal care for women who have become pregnant from rape; fistula surgery; interventions to restore fertility; interventions to repair damage to sexual organs; etc. Appropriate medical services will vary from context to context. The indicator should be coded as a “0” if no specialized medical services are offered to SGBV victims, or if only general health care is offered. For example, if SGBV victims are simply enrolled in the national health care service, then the indicator should be scored a “0.”

Satisfaction: The reparations programme has implemented at least one symbolic measure specifically targeted to women and/or girls. In countries where SGBV and gender discrimination is deeply entrenched, symbolic measures such as apologies and memorials can be important in signaling the state’s commitment to broader change toward gender justice. For example, in 2010, the President of Sierra Leone formally apologized to women victims of the country’s conflict; in Kenya, the recent report of the Truth, Justice and Reconciliation Commission recommended a similar apology from the head of state, although it has not yet taken place. As mentioned, symbolic measures include official apologies and commemorations and tributes to victims, such as public memorials, holidays, or other officially sponsored acts. In order to be scored as a “2” or a “1,” the measures must be “specifically targeted to women/girls.” Such targeting can take many forms. Examples include official apologies to women/girls for violence they experienced during conflict and

commemorative days or public memorials that honor women victims.

- **Non-repetition: The state has adopted and implemented at least one legal reform intended to promote gender equality, stemming from the recommendations of a truth commission or from an official reparations policy.** The *Basic Principles and Guidelines* insist that the state has a responsibility to take steps to prevent the recurrence of human rights abuse. For women and girls in many countries, legal reforms are necessary to guarantee rights that act as bulwarks against SGBV, dispossession, and poverty. Examples include the following reforms: property and/or matrimonial property laws; inheritance laws; marriage laws; laws on sexual violence; child protection laws; and equal opportunity laws. Such reforms are often included as recommendations in the final reports of truth commissions. In order to be scored as a “2,” the state must not only adopt, but also put in place adequate administrative and budgetary measures to implement the law. Cases in which a new law exists, but where there is a lack of resources or administrative capacity to implement it, should be scored a “1” until such time as the needed resources and capacity are strengthened. Cases where there have been legal reforms to promote gender equality, but the reforms are not clearly linked to a reparations process, should be scored a “0.”

The target for the basket is 100 percent compliance with all five indicators.

Rationale

The new approach offers a number of advantages. The conceptual relevance is more clearly related to the objective, which is to assess states’ commitments to the effective inclusion of women. Moreover, the basket approach directly addresses the challenges to women’s

effective inclusion in reparations programs that the UN has identified, such as the lack of inclusion of SGBV as a reparable harm. Finally, this approach is tied to the five types of reparations outlined in the *Basic Principles and Guidelines*. Each indicator is designed to signal one gender-related dimension of each type of reparations. It should be noted that the indicators are not intended to give a comprehensive picture of each of the five types of reparations, but rather provide a targeted signal that can be useful in assessing whether reparations programs are generally heading in a positive direction.

Moreover, data collection for this indicator will be easier. The basket primarily consists of a series of yes/no questions about specific benefits that have actually been implemented. The questions can be clearly answered by expert observers, and confirmed independently. It thus avoids the many thorny issues involved with data collection and measurement, especially for those reparations programs that distribute benefits through a range of local and state agencies.

Reflections and Lessons Learned

The basket indicator ultimately developed for UN Women was a compromise solution to a very interesting puzzle: can you develop a single indicator that is a reliable signal of a country's performance on a TJ issue? The answer, for me, is no. Although I have described a very particular example in this essay—an indicator on reparations for UN Women—many of the challenges that the example raises are generally applicable to the TJ field as a whole. Ultimately, only a set of indicators that embrace a range of dimensions of TJ would really be useful to policy and program actors—and indicators are supposed to be **useful to** someone. Discussions with colleagues at UN Women, the International Center for Transitional Justice, the International Organization for Migration, and elsewhere suggest a number of gaps that would need to be filled in order for international actors to create cross-national indicators of the kind that development actors currently use.

First, TJ initiatives themselves should be producing reliable, public administrative data—not just on budgets and expenditures, but also on outputs and outcomes. This is important not just for the creation of indicators for international actors, but perhaps even more so that national and local civil society groups can monitor TJ processes, applying pressure and advocacy when necessary. The data needs to be gathered so that it can be disaggregated across a range of criteria, including gender, ethnicity, region, and income level.

Second (and related), TJ actors need to develop capacity on data production, or partner with impartial, expert agencies that can provide the needed technical expertise. In many countries, the national statistical body may be a possible source of support; moreover, engaging this body may open a productive dialogue on how to measure human rights—a field with which those with statistical training may be unfamiliar. Elsewhere, we have seen TJ institutions, especially truth commissions, partner with international NGOs with expertise in database creation, which can also be a useful approach. As an aside, I found this process with UN Women to be a positive example of mutual learning between the “program” and the “statistics” side of UN Women. Each side pushed the other to recognize legitimate needs as well as one’s own boundaries of expertise. On the program side, it was important to emphasize the limited nature and role of indicators—they cannot and should not tell us everything. It’s just not feasible. On the statistics side, I believe that there was a stronger appreciation for the differences between the objectives of “development” and “human rights” indicators.

Third, it would be useful for data to be centralized as much as possible, especially since the implementation of TJ is almost always distributed across different governmental bodies and levels. This ideal is probably not achievable in most cases—except, perhaps, in countries like Tunisia that have set up specific ministries or agencies for dealing with TJ. At a minimum, it would be useful for the various agencies working on TJ to be in contact with one another in order to set basic

parameters for data development, so that similar data is being collected.

Finally, the utility of indicators is that they are signals of change—defining “which” changes are the most important ones to track is critical to the development of cross-national indicators. UN norms provide guidance—albeit highly complicated and evolving—on the most important changes to seek. However, this thicket of documents needs to be simplified in terms of specific and measurable indicators upon which people can generally agree—not just donors and multilateral agencies, but also the countries where TJ is currently operating. So far, this very important task has not yet been done. UN Women has taken the first step, but it is not yet clear that there will be a second one, from UN Women or any other UN agency. It might be easier for policy actors to push ahead with indicator development if there were more readily available data to work with. It may be, however, that states need to be incentivized to create this data; having UN-created indicators that states must report on is one way to push countries to develop data in the first place. At a minimum, those countries that are receiving financial or technical assistance from the UN on TJ issues should be pushed to produce more, better, and publicly available data on their design, implementation, and outcomes.

Appendix

Proposed scoring tool: Indicator 26a on Reparations

Indicator	Score 0, 1, 2
The reparations programme has implemented cash payments and/or pensions specifically for widows/spouses	
A restitution programme exists, and it receives a score of “2a” or above according to IASC Gender Marker codes (i.e., the project is designed to contribute significantly to gender equality)	
The reparations programme has implemented medical services specifically for victims of SGBV	
The reparations programme has implemented at least one symbolic measure specifically targeted to women and/or girls	
The state has adopted and implemented at least one legal reform intended to promote gender equality, stemming from the recommendations of a truth commission or from an official reparations policy	
TOTAL BASKET SCORE (out of 10)	

Explanation of Scoring

Each indicator should be assessed a score of “2” if it has been fulfilled, “1” if it has been partially fulfilled (i.e., is still ongoing), and a score of “0” if it has not been fulfilled at all. A short, one-paragraph description should be included to give the rationale for the score for each of the five indicators. This description should include the sources of information for the scoring, as well as:

In the case of a score of “2”:

- What steps were taken to fulfill the indicator, including specific information on implementation
- When the indicator was fulfilled

In the case of a score of “1”:

- What steps have been taken so far
- An estimate of when the indicator will be fulfilled in the future
- In the case that there has been no change since the prior assessment of the indicator at the “1” level, an explanation for why there has been no change

In the case of a score of “0”:

- Why no steps have been taken to fulfill the indicator
- In the case that there has been no change since the prior assessment of the indicator at the “0” level, an explanation for why there has been no change